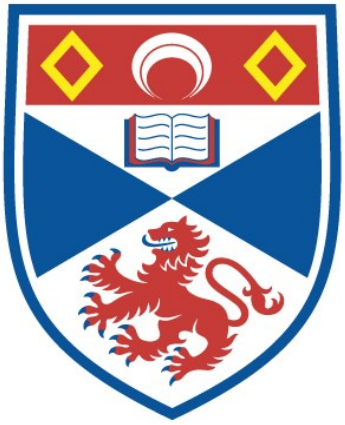
MOLECULAR ASPECTS OF NITRATE ASSIMILATION IN 'ASPERGILLUS NIDULANS'

Kim Lorraine Hawker

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



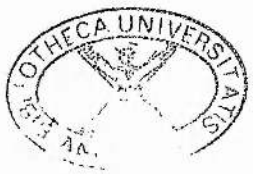
1991

Molecular Aspects of Nitrate Assimilation in *Aspergillus nidulans*

Kim L. Hawker

Ph.D. Thesis

1990



To my parents

Declaration for the degree of Ph.D.

I, Kim Lorraine Hawker, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed..... Date....6/12/90....

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 1st October, 1987 and as a candidate for the degree of Ph.D. on 1st November, 1988.

Signed..... Date....6/12/90....

Certificate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree or Ph.D.

Signature of Supervisors:

..... Date 15/12/90

..... Date 2/12/90

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bone fide* library or research worker.

Acknowledgments

I would like to thank my supervisor Dr. Jim Kinghorn for his continual guidance throughout the course of this project. I am grateful for the support and encouragement provided by Dr. Paul Montague, to whom I am also indebted for the demonstration of unfamiliar, molecular techniques. In addition, I wish to express my gratitude to my colleagues and fellow students for advice and assistance. In particular, I convey thanks to Dr. Shiela Unkles, Dr. Michael Abberton, Miss Carolyn Grieve and Mr. Eddie Campbell and to Dave and Sean for the excellent photographic work. Finally, I wish to thank all those who contributed towards making my stay in St. Andrews enjoyable. This work was supported by an SERC quota award.

ABSTRACT

The *crnA* gene of *A.nidulans* is confirmed to encode a transport protein, possibly for nitrate. The single open reading frame of 1449 nt encodes a polypeptide of 483 AAs with a molecular weight of 51,769 Da. The high percentage of clustered apolar AAs suggests a membrane protein with ten membrane spanning domains. No similarity between the *crnA* encoded polypeptide and those of other transport genes or any other gene was observed. The regulation of the *A.nidulans* nitrate transporter was investigated by studying levels of mRNA transcribed from the *crnA* gene. The *crnA* cDNA hybridised to two mRNAs on Northern blots. The 1.8 kb message corresponding to the size of the *crnA* gene is inducible with nitrate and nitrite. The 1.1 kb, constitutively synthesized message is believed to be a second transcript from the *crnA* gene rather than from a gene elsewhere in the genome. The expression of the *crnA* gene, determined from Northern blotting experiments, is shown to be at the level of mRNA accumulation. A requirement for both positive acting regulatory genes, *areA* and *nirA* is observed.

The expression of the *niiA* and *niaD* genes, encoding nitrite reductase and nitrate reductase respectively, is shown to be regulated at the level of mRNA accumulation by a mechanism involving both the *areA* and *nirA* gene products. In addition, the nitrate reductase apoenzyme influences the expression of all three structural genes *crnA*, *niiA* and *niaD*. This effect is also exerted at the level of mRNA accumulation.

Sequences upstream of the *niaD* and *niiA* genes from *Aspergillus oryzae*, *A.nidulans* and *A.niger* were compared, in addition to a similar

region of the *crnA* gene. In particular, one sequence of ten nucleotides was found to share substantial similarity between each gene. It is possible that this motif is a *cis*-acting transcription factor recognised by a regulatory protein. The consensus TATCTA, a possible receptor site for the *areA* gene product, was identified several times upstream of each gene.

The *nit-4* gene of *Neurospora crassa* is shown to complement the *A.nidulans nirA* loss of function mutant via genetic transformation. The three transformant strains each have multiple integrates of the *nit-4* gene, although none are present at the homologous site. The regulation of *A.nidulans* nitrate reductase by the *N.crassa nit-4* gene is observed to be essentially normal, i.e. wild type.

Abbreviations used in the text

i) general

α	alpha
A	adenine
AA	amino acid
ATP	adenosine triphosphate
ß	beta
bp	base pair
C	cytosine
°C	degrees centigrade
cDNA	complementary DNA
cm	centimetre
CM	complete media
CTP	cytidine triphosphate
δ	delta
Da	Dalton
DEP	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dpm	disintegrations per minute
EDTA	ethylenediamine tetra acetic acid
ER	endoplasmic reticulum
et al	et alia (and others)
FAD	flavin adenine dinucleotide
Fig.	figure
γ	gamma
G	guanine
GTP	guanosine triphosphate
g	gram
h	hour(s)
hnRNA	heterogeneous nuclear RNA
hnRNP	heterogeneous nuclear RNP
k	1 thousand
kb	kilobase
kDa	kiloDalton
kg	kilogram
λ	lambda
L	litre
LB	Luria broth
μg	microgram
μl	microlitre
μm	micrometre

µM	microMolar
M	Molar
MM	minimal media
mg	milligram
min	minute(s)
ml	millilitre
mM	milliMolar
mm	millimetre
Mr	relative molecular weight
mRNA	messenger RNA
N	any nucleotide
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
ng	nanogram
NiR	nitrite reductase
nm	nanometre
nM	nanoMolar
No.	number
NR	nitrate reductase
nt	nucleotide(s)
OD	optical density
pers. comm.	personal communication
pg	picograms
pM	picoMolar
R	any purine
RNA	ribonucleic acid
RNase	ribonuclease
RNA pol II	RNA polymerase II
rpm	revolutions per minute
rRNA	ribosomal RNA
s	seconds
snRNA	small nuclear RNA
T	thymine
TFIID	transcription factor II D
Tris	Tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
TTP	thymidine triphosphate
U	units
U	uracil
UAS	upstream activator sequence
UV	ultraviolet
V	volts
v/v	volume per volume
WT	*A.nidulans* wild type strain
w/v	weight per volume
X	any amino acid
Y	any pyrimidine

ii) amino acids

	three letter code	one letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

List of figures

List of tables

List of plates

Contents

Chapter 1

INTRODUCTION

1.1.0 BACKGROUND

Concepts of biosynthetic pathways and gene : enzyme relationships have, in many cases, been developed from genetic and biochemical studies of filamentous fungi. In particular, much work has been focussed on the enzymes and genes involved with nitrate assimilation in *Aspergillus nidulans* (for a review see Cove, 1979). The two structural genes, *niaD* and *niiA*, encoding the enzymes nitrate reductase (NR) and nitrite reductase (NiR) respectively (Cove and Pateman, 1963 and Pateman *et al*, 1967), were found to be positively regulated by a pathway specific gene, *nirA*, the product of which is required for induction (Cove and Pateman, 1969) and a second gene, *areA*, with a wider regulatory role, encoding a protein responsible for nitrogen metabolite repression (Cohen, 1972, Arst and Cove, 1973 and Hynes, 1975). Although the kinetics of nitrate uptake have been briefly analysed (Brownlee and Arst, 1983) the gene, *crnA*, which is thought to encode a nitrate permease requires a more detailed examination and is a major focus of this thesis.

An analysis of the structure and regulation of a eukaryotic nitrate permease is particularly important since as yet, no other has been investigated and until recently no information on any anionic transporter was available (Mann *et al*, 1989). Along with this, further analysis of the structural genes involved in nitrate assimilation, i.e. *niaD* and *niiA*, is appropriate due to the increased use of nitrates in agriculture and their subsequent effects on the environment, for example eutrophication. Hence, the importance of attempting to extend our understanding of nitrate assimilation.

Filamentous fungi offer many of the genetic complexities of a eukaryote with the ease of manipulation of a micro-organism (Rambosek and Leach, 1987). The genus *Aspergillus* is included within the fungal class of Ascomycetes and consists of a number of species that have been

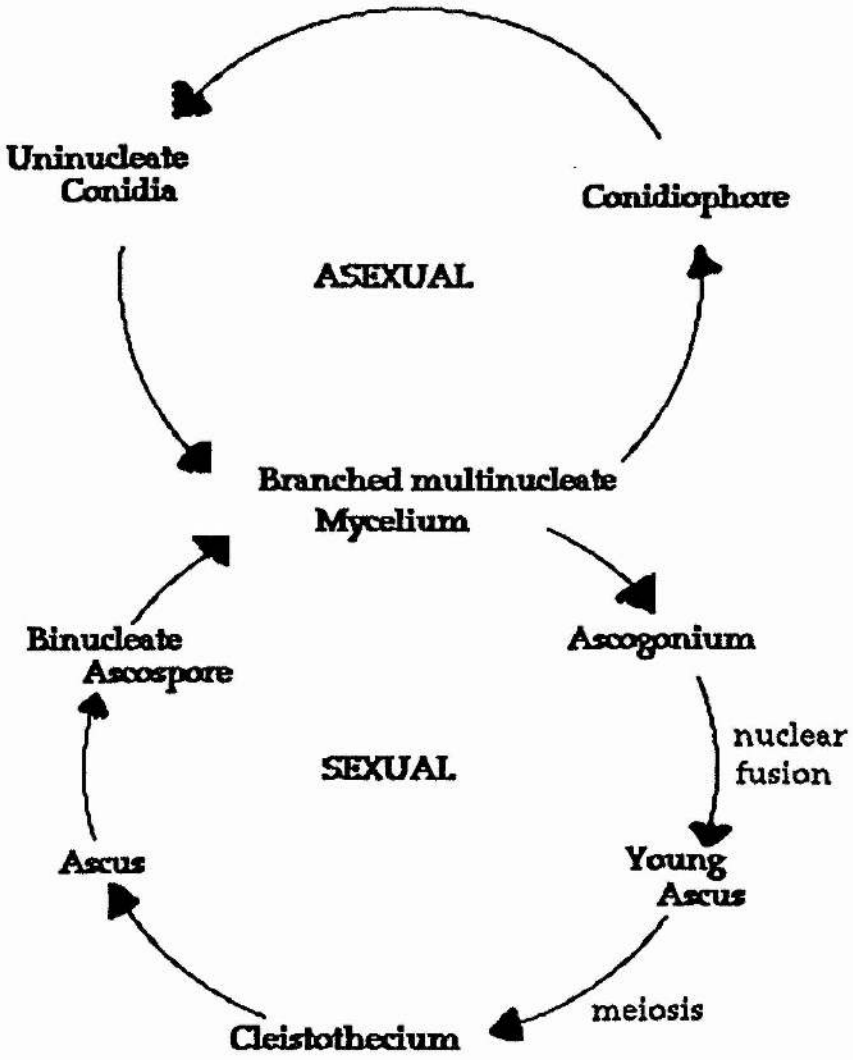


Figure 1.1 Life cycle of *A.nidulans* showing both sexual and asexual cycles. The parasexual cycle is missing.

investigated at the biochemical and genetic level. *Aspergillus oryzae* and *Aspergillus niger*, important commercially for the production of primary metabolites and enzymes, for example citric acid (Kapoor *et al*, 1982) and α-amylase (Boing, 1982), are not as well characterised as *A.nidulans* since this species, in particular, has a sexual life cycle aiding genetic manipulation (Fig 1.1). Other general advantages offered by *A.nidulans*, are reviewed by Pontecorvo *et al* (1953). The asexual progeny, uninucleate conidia, grow rapidly as a single colony with minimal nutritional requirements. In addition, *A.nidulans* has metabolic versatility, homothallism, multi-nucleate hyphae and both haploid and diploid growth phases. The position of many genes on the eight chromosomes have been examined providing a fairly detailed genetic map (Clutterbuck, 1984) and several of these genes have now been physically characterised (1.2.1).

A.nidulans is therefore, a suitable organism to use as a model when investigating gene structure and regulation. In addition, the relatively small genome size of 2.6×10^7 bp with low repetitive DNA content (Timberlake, 1978), offers advantages over mammals and plants in techniques involving modern DNA technology, particularly of library construction and screening. The availability of a variety of metabolic mutant strains, DNA mediated transformation and gene cloning technology may well lead to a better understanding of gene regulation.

1. 2. 0. EUKARYOTIC GENE STRUCTURE

The birth of modern molecular biology techniques has led to the isolation and structural analysis of a large number of genes from phylogenetically diverse organisms. Analysis at the nucleotide level has provided information on the structure and regulation of both prokaryotic

and eukaryotic genes. There are several features that distinguish eukaryotic genes from prokaryotic genes. First, eukaryotic genes are found on several chromosomes rather than just the one circular chromosome of prokaryotes (excluding those bacteria with plasmids).

Second, unlike prokaryotic DNA, the vast majority of eukaryotic DNA is bound to histones and numerous non-histone proteins, forming chromatin (Weisbrod, 1982). The basic unit of chromatin is the nucleosome consisting of 146-240 bp of DNA wrapped twice around a histone core. This is made up of two of each of the four inner histones H2A, H2B, H3 and H4. Between 10 and 100 bp of DNA separates each nucleosome. Histone H1 is present at one copy per nucleosome and is possibly also associated with the linker DNA. Nucleosomes are further wound into higher order structures ultimately reducing the DNA to a one thousandth of its original length (Weisbrod, 1982).

As the genome size increases in eukaryotes so too does the quantity of non-coding sequences. The presence of these sequences may impose a greater energy requirement for DNA replication, an unnecessary burden for lower eukaryotes, having a higher replication rate. Alternatively, these sequences may represent selfish genes which have evolved only for survival within the genome, rather than for expression (Doolittle and Sapienza, 1980).

Eukaryotes have three different types of RNA polymerase, the enzyme complex responsible for transcription. Each transcribes a different set of genes. Those genes encoding large rRNAs belong to class I, class II includes genes that encode proteins and some snRNAs whilst class III genes are those encoding tRNAs and the remainder of the snRNAs.

The exons of many eukaryotic genes are interrupted by non-coding regions, termed introns. Exons represent the protein coding sequences of a gene and separation from introns could sort protein functions to

produce novel proteins. This is thought to increase the rate of evolution (Blake, 1983).

Finally, the presence of a nuclear membrane in eukaryotes separates the processes of transciption of DNA to RNA and translation of RNA to polypeptide, both in space and time.

1. 2. 1. An overview of filamentous fungal pol II gene structure

The genes under study in this thesis encode proteins transcribed by RNA pol II. It is difficult to generalise on the structure of pol II genes since there is great diversity, especially between organisms from different phylogenetic orders. However, features common to several filamentous fungal genes are dicussed (Fig 1.2).

The position and complexity of the transcriptional start site in filamentous fungal genes varies greatly. The tendency is for several sites usually with one predominating. For example, the *qa-1S* gene (Huiet and Giles, 1986) and the *am* gene (Kinnaird and Fincham, 1983) of *Neurospora crassa* and the *oliC* gene of *A.nidulans* (Ward and Turner, 1986). Many genes have only one transcriptional start site, including the *amdS* gene of *A.nidulans* (Corrick *et al*, 1987) and the gene encoding the ADP/ATP carrier of *N.crassa* (Arends and Sebald, 1984). However, these observations may reflect the sensitivity of the method used for 5' end determination. The length of the leader sequence of filamentous fungal genes, i.e. the distance between the transcriptional and translational start sites, may be as great as 400 nt (the *qa-4* gene of *N. crassa* (Rutledge, 1984)) or as short as 28 nt (the *trp-1* gene of *N.crassa* (Schechtman and Yanofsky, 1983)). A structure known as the cap is added to the 5' end of the primary transcript, generally before the hnRNA chain is longer than 50 nt. Its purpose is to enhance the stability of the mRNA in the cytoplasm.

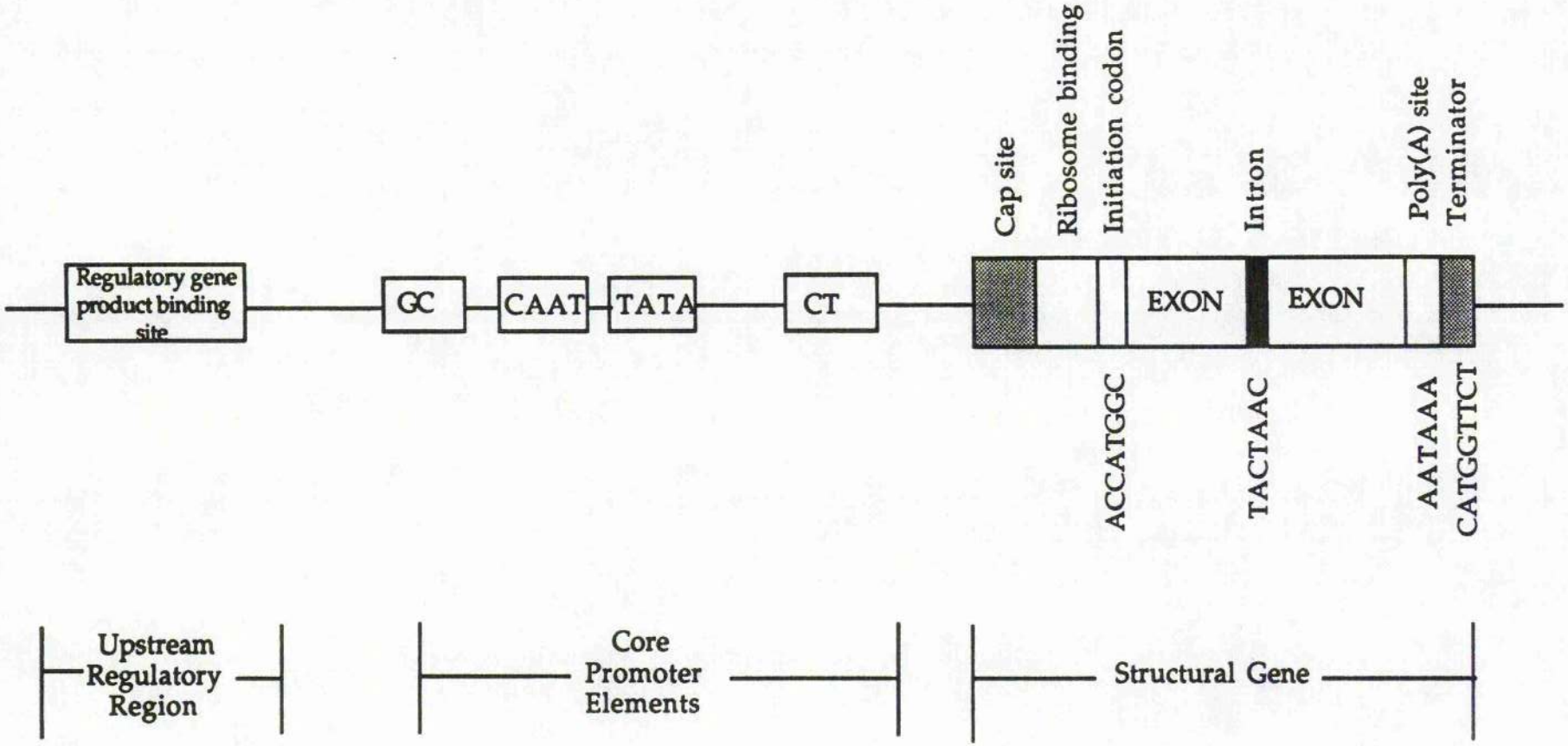


Figure 1.2 The structure of a "typical" filamentous fungal pol II gene. Some features included are more commonly found in mammalian genes especially the GC and CAAT boxes and the poly(A) addition site.

Capping involves a 5' - 5' pyrophosphate linkage between 7-methyl-guanosine and the hnRNA (Coppola *et al*, 1983).

The start of the coding sequences is usually at the first ATG after the trancriptional start site. In mammalian genes it has been observed that the consensus of ACCATGGC surrounding the ATG site is the optimal sequence for ribosome binding, necessary for the initiation of translation (Kozak, 1986). A similar sequence can be found in most fungal genes.

The exons of mammalian genes are generally short, interrupted by many long introns, contrasting sharply to the genes of filamentous fungi. In the majority of these genes sequenced to date, introns are few; occuring either once, as in the case of the *his3* gene of *N.crassa* (Legerton and Yanofsky, 1985) and the *prnB* gene of *A.nidulans* (Sophianopoulou and Scazzocchio, 1989) or twice, for example the *am* gene of *N.crassa* (Kinnaird and Fincham, 1983) and the *aldA* gene of *A.nidulans* (Pickett *et al*, 1987). Introns may be absent from some genes whereas the *gpdA* gene of *A.nidulans* has seven introns (Punt *et al*, 1988) and the *niaD* and *niiA* genes have six and seven introns respectively (Johnstone *et al*, 1990). With the exception of several introns within the *actA* gene of *A.nidulans* (Fidel *et al*, 1988), the length of filamentous fungal introns is generally less than 100 bp.

Introns are less common in yeast genes, although the consensus sequences AG'GTATGT and Py(11)NCAG'GT are found at the exon : intron and intron : exon boundaries respectively, with splicing occurring between the two guanidine residues in both cases (Langfordand Gallwitz, 1983). The consensus PyGCTAACN, known as a branch-point sequence, found near the 3' end of the intron is involved with the formation of a lariat during splicing (Langford and Gallwitz, 1983 and Sharp, 1985), whereby the 5' end of the intron is esterified with an adenosine residue

in the conserved sequence. The two exons are subsequently ligated together after the circularised intron is released at the 3' splice site. Similar sequences may be found in filamentous fungal introns. That filamentous fungal introns are not identical to yeast introns is consistent with yeast being unable to excise the intervening sequences from foreign genes (Langford *et al*, 1983).

In higher eukaryotic genes the branch-point sequence is not highly conserved. Cryptic branch-points are activated when the authentic sequence is mutated, the location of 18-37 nucleotides from the 3' end of the intron being the important factor in determining the location of lariat formation (Green, 1986). Therefore, although splicing is highly conserved in mechanism and use of homologous signals it is executed by a machinery which is variable at the species level. In addition, it would appear that yeast, possessing few genes with introns, does not have a flexible machinery to cope with sequence degeneracy as is the case with, particularly higher eukaryotes and some filamentous fungal systems.

It has been observed that codon usage in highly expressed filamentous fungal genes is more biased than those genes with a low level of expression (Punt *et al*, 1988 and Clements and Roberts, 1986). In particular, there is strong preference for codons ending in a pyrimidine nucleotide, where more than one codes for a single amino acid. Codon usage in the *gpdA* gene of *A.nidulans* is 93% biased towards a pyrimidine residue in the third position, when a choice is allowed (Punt *et al*, 1988). Such a bias is not observed with the codons of *S.cerevisiae* (Bennetzen and Hall, 1982). Codon utilisation has been considered as a possible regulator of gene expression, although it is more likely that codon utilisation is modulated for gene expression (Holm, 1986).

The protein coding region of genes terminates at a stop codon, UAA, UAG or UGA. Downstream, the consensus sequences AATAAA

and CAYTG (Berget, 1984) may occur, mediating poly(A) addition which, like the cap, is thought to enhance the stability of the mRNA. Sequences similar to the AATAAA consensus have been shown to allow polyadenylation, albeit at a reduced efficiency (Wickens, 1990). Two additional regions, GT and T-rich sequences, are also implicated in polyadenylation, the efficiency decreasing as the distance between the two sequences increases (Gil and Proudfoot, 1987).

Initial sequence comparisons of yeast genes proved the consensus TAAAATAAG to mediate polyadenylation (Bennetzen and Hall, 1982). However, it was later suggested that the occurrence of this consensus is a coincidence as many 3' non-coding regions of yeast genes are AT-rich. Instead the motifs TAGT and TAGTGT were proposed to be necessary for polyadenlyation (Zaret and Sherman, 1982). Characterisation of the homoserine O-transacetylase gene, *met-2*, of the fungus *Ascobolus immersus* (Goyon *et al*, 1988) identified a sequence (T-rich)---TAG---TAGT---TTT downstream of the coding region thought to be involved with poly(A) addition.

Recently, polyadenylation and 3' end formation in yeast have been shown to be independent of the AATAAA motif, supporting the above observation. In addition, the protein fraction of yeast responsible for 3' end formation does not recognise the human α-globin AATAAA signal, suggesting that different poly(A) machinery exist in yeast and mammalian systems (Butler *et al*, 1990). It is proposed that a similar mechanism to that of yeast may operate for polyadenylation in filamentous fungi since few fungal genes have the AATAAA motif (exceptions include the *trpC* gene of *A.nidulans* (Mullaney *et al*, 1985)).

Termination of transcription usually occurs 150-200 bp downstream of the poly(A) tail. The consensus YGTGTTYY, found in

many mammalian genes, may be a termination signal (McLauchlan *et al*, 1985).

The *cis*-acting elements are sequences within the 5' non-coding regions of genes that are proposed to interact with *trans*-acting factors modulating the rate of transcription. Such elements may be promoter-specific sequences or elements responsible for binding the products of regulatory genes; the upstream activator and repressor sequences of yeast and the enhancer and silencers of mammalian systems.

TATA, CAAT and GC motifs may be found in the promoter region of filamentous fungal genes, along with a TC-rich region. They may occur uniquely or in conjunction with one or more of the remaining. The GC box is more common in mammalian genes (1.3.1.) although it has been identified in the *alcA* (Gwynne *et al*, 1987) and *aldA* (Pickett *et al*, 1987) genes of *A.nidulans*. The CAAT box (Breathnach and Chambon, 1981) is found 60-120 bp upstream from the transcriptional start site of several filamentous fungal genes including the *qutE* gene of *A.nidulans* (Da Silva *et al*, 1986) and the *pyr-4* gene of *N.crassa* (Newbury *et al*, 1986). However, it too is more generally associated with mammalian genes and its functional significance in filamentous fungi remains to be determined.

A sequence comparison of 41 mammalian genes identified a common TATA motif in a similar position within the 5' non-coding sequences (Corden *et al*, 1980). Such a motif, indispensible for the initiation of transcription, is present 30-70 bp upstream of the transcriptional start site in the majority of fungal genes sequenced to date. The sequence may be duplicated, as in the sesquiterpene cyclase gene of *Fusarium sporotrichioides* (Hohn and Beremand, 1989) or present in multiple copies, for example the calmodulin gene of *Achlya klebsiana* (LeJohn, 1989).

One sequence not found in mammalian genes that occurs frequently in yeast and filamentous fungal promoters is a pyrimidine-rich region, i.e. the CT block. (Examples include the cutinase gene of *Fusarium solani* (Soliday *et al*, 1989) and the *gpdA* gene of *A.nidulans* (Punt *et al*, 1988)). The length of this region, situated between the TATA box and the transcriptional start site, has been proposed to be directly related to the level of transcription (Ballance, 1986). However, functional analysis of CT boxes in the *S.cerevisiae CYC1* gene has indicated that they may play a role in fixing the transcriptional start site (McNeil and Smith, 1985).

1. 2. 2. Determination of putative upstream *cis*-elements

Potential *cis*-acting sequences, receptor sites for the products of activator genes, have been identified in many systems initially by computer-aided comparisons of upstream non-coding regions of related genes. Comparisons of the 5' flanking regions of the *alcA* and *aldA* genes of *A.nidulans*, involved with ethanol utilisation, revealed six similar sequences of between 4 and 24 nt in length. Each region occupies a similar relative position upstream from the translational start site of both genes. The motif furthest from the ATG shares substantial identity with the promoter of the yeast *ADH2* gene, also involved with ethanol utilisation. This suggests that the product of the *alcR* regulatory gene may interact at this site, activating the *alcA* and *aldA* genes (Gwynne *et al*, 1987).

A single sequence of consensus GGNTAARYRYTYATCC is repeated 13 times within the 5' non-coding sequences of the *qa* gene cluster, involved with quinic acid utilisation in *N.crassa*. The sequence has partial dyad symmetry and has been shown, by DNA-binding and DNase I foot-printing, to bind the product of the *qa-1F* activator gene. That multiple sequences occur upstream of some genes is proposed to

allow for greater transcriptional control by the *qa-1F* gene product (Baum *et al*, 1987). The upstream regions of the *qut* genes for quinic acid utilisation in *A.nidulans* share a 16 nt consensus sequence GCCAGANCGTTCTNCC, similar to that of the *N.crassa qa* gene cluster. In addition, a 9 nt motif CGACAATCT is found on the non-translated strand upstream of each *qut* gene (Hawkins *et al*, 1988).

The short sequence TGACTC is repeated seven times within the 5' flanking regions of the *argB* gene, for arginine utilisation, in *A.nidulans*. Genes of *S.cerevisiae* under general amino acid control also have the sequence TGACTC in their upsteam regions (Goc and Weglenski, 1988). In addition, the GCN4 transcription factor of yeast binds general control promoters at all TGACTC sequences (Arndt and Fink, 1986) suggesting that this motif may be promoter-specific rather than a *cis*-element involved with the regulation of only a few, related genes.

The 5' flanking sequences of the *A.nidulans gpdA* gene, encoding glyceraldehyde-3-phosphate dehydrogenase were compared with similar regions of other filamentous fungal genes. Sequences sharing identity with the *qa* and *qut* consensus were identified. A comparison of the upstream regions of the *A.nidulans* and *A.niger gpdA* genes revealed a homologous 50 nt region occupying a similar position from the ATG codon of both genes. Fusion to the *E.coli lacZ* gene (4.7.0) confirmed this region to be involved with transcription factor- or regulatory protein-binding (Punt *et al*, 1990).

The product of the *N.crassa nit-2* regulatory gene for nitrogen metabolite repression recognises elements containing at least two copies of the consensus TATCTA uptream of the *nit-3* gene, encoding nitrate reductase. Four such areas within the upstream sequences of the *A.nidulans niaD* gene are also capable of binding the *nit-2* gene product (Fu and Marzluf, 1990). This supports previous results (1.4.7), providing

more than one site may be required for successful binding of regulatory gene products.

Computer-aided comparisons of 5' flanking regions of genes have therefore successfully identified *cis*-elements in a number of systems. From the information provided here it would appear that these activator sequences may consist of direct or indirect repeats, sequences capable of forming hairpin-loops or simply a sequence without any particular pattern. Silencers are also *cis*-acting sequences with which the products of regulatory genes interact. However in contrast to those *cis*-elements discussed above the silencer is responsible for deactivating gene expression.

1. 3. 0. THE REGULATION OF POL II GENES

The regulation of gene expression is exerted at a variety of levels: transcriptional control, hnRNA processing, transport of mRNA across the nuclear membrane, mRNA stability, translation and post-translation (Darnell, 1982).

1. 3. 1. Transcriptional control

The modulation of hnRNA chain initiation is determined by several inter-related factors, for example the chromatin structure, *cis*-elements and *trans*-acting proteins.

Experiments with mammalian systems, *Xenopus*, *Drosophila* and *S.cerevisiae* demonstrate that the structure of the chromatin plays a crucial role in gene expression (Weisbrod, 1982). For instance, histone H1 is responsible for organising the higher order structure of the nucleosomes (1.2.0) Adjacent nucleosomes within inactive regions of chromatin are held together by histone H1 preventing the transcription

nucleosomes (1.2.0) Adjacent nucleosomes within inactive regions of chromatin are held together by histone H1 preventing the transcription factors from gaining access (Weintraub, 1985). DNase I studies demonstrate that *cis*-acting regions of genes become histone free during transcription (Mills *et al*, 1983).

Left-handed, Z-DNA is a conformational switch involved with the control of transcription. Unstable under normal physiological conditions it is stabilised by either supercoiling, binding to proteins specific for Z-DNA, binding specific ions or by methylation (Nordheim *et al*, 1981). Topoisomerase II induces and maintains negative supercoiling, altering the chromatin structure, a feature of active genes. Proteins with an affinity for superhelical DNA can then bind, initiating transcription (Villeponteau *et al*, 1984).

The *cis*-elements are DNA sequences responsible for binding specific factors that ultimately either increase or decrease transcription (1.2.0). The GC rich motif, with the consensus GGGCGG or its complement, is prominent in mammalian house-keeping genes, but unidentified in microbes and plants. The transcription factor Sp1 binds to the GC box (Dynan *et al*, 1985) which is considered to maintain the gene at a constitutive level of activity (Dynan and Tjian, 1985). Several GC boxes may be present in a promoter, the site closest to the gene usually mediating the strongest stimulus (Jones and Tjian, 1985).

The CAAT box is capable of binding a multiplicity of factors, including CTF and NF-1 (Dorn *et al*, 1987), important in mammalian systems for maintaining low level or constitutive expression.

The TATA box, the most ubiquitious conserved signal throughout the eukaryotic kingdom, may be important for the regulation of access of RNA pol II (Darnell, 1982), which recognises changes in the chromatin structure (Baer and Rhoades, 1983). Recently, five protein factors have

been recognised that complex with the RNA pol II and the DNA resulting in a basal level of transciption. These are the general transciption factors II A, B, D, E and F. TFIID is the only factor capable of direct interaction with the DNA, the others forming a complex with the RNA pol II (Mermelstein *et al*, 1989).

In addition to these promoter-specific sequences, other *cis* regions of the DNA exert an effect on transcription. Upstream activator sequences (UAS) and upstream repressor sequences are two regions common to many yeast pol II genes which modulate the rate of transcription. The protein factor HAP1 of *S.cerevisiae* binds to the UAS2 of the *CYC1* and *CYC7* genes involved in the production of cytochrome c oxidase. In this way the level of transcription of these genes is increased (Pfeifer *et al*, 1987).

Enhancers are considered to be the higher eukaryotic equivalent of the yeast UASs. There is a large variation in their size, ranging from 40 to 700 bp, comprising of many short, repeated sequences. The effect enhancers exert on transcription may be from as far away as 10 kb from the cap site (Pinkert *et al*, 1987) or at the 3' end past the poly(A) site (Campo *et al*, 1983). In addition, an immunoglobulin gene posesses an enhancer within an intron (Mercola *et al*, 1983) and the enhancer of the hepatitis B virus is within a protein-coding region (Tognoni *et al*, 1985). It follows that enhancers act in either orientation exerting regional transcriptional control for functions encoded by either DNA strand (Khoury and Gruss, 1983).

The effects on regional transcriptional activation of enhancer and UAS elements may be to either induce changes in the DNA conformation (Villeponteau *et al*, 1984) or the chromatin structure (Mills *et al*, 1983). *Trans*-acting proteins may bind to these regulatory elements (Sassone-Corsi and Borrelli, 1986), subsequently making contact with the

promoter-binding factors (Ptashne, 1986). However, their importance may only be in the establishment and not the maintenance of the transcriptional complex (Wang and Calame, 1986). It is believed that combinations of *cis*-elements arranged in a unique configuration confer on each gene an individualised spatial and temporal transcriptional programme (Mitchell and Tjian, 1989).

A number of regulatory proteins encoded by RNA pol II genes have been shown to be responsible for interacting with the *cis*-elements and either initiating or repressing transcription (1.2.2). The DNA binding domains of these proteins are usually at the N-terminus often characterised by any of the following five structures: the helix-turn-helix, the two types of zinc fingers, leucine zippers and the helix-loop-helix (for a review see Glover, 1989) (Fig 1.3). The C-terminus, rich in acidic ("acid blobs"), proline or glutamine residues, is directly responsible for transcriptional activation by protein : protein contacts with factors bound to promoter regions (Mitchell and Tjian, 1989 and Sigler, 1988).

The helix-turn-helix (Fig 1.3) is common to prokaryotic activator and repressor proteins. It includes two helices separated by a turn within the protein structure. The second helix is responsible for most of the sequence-specific contacts with the DNA and is referred to as the recognition helix (Zhang *et al*, 1987). In eukaryotes, a similar sequence to the helix-turn-helix was first found in a family of *Drosophila* proteins that control early development. The 180 bp highly conserved sequence, known as the homeobox encodes a 60 AA homeodomain common to yeast and man (Struhl, 1989).

Zinc fingers (Fig 1.3) are loop-like structures surrounding a zinc atom that interact with specific regions of DNA. Although there is diversity in the nature and spacing of the liganding residues and the exact sequence surrounding them, zinc fingers can be divided into two types.

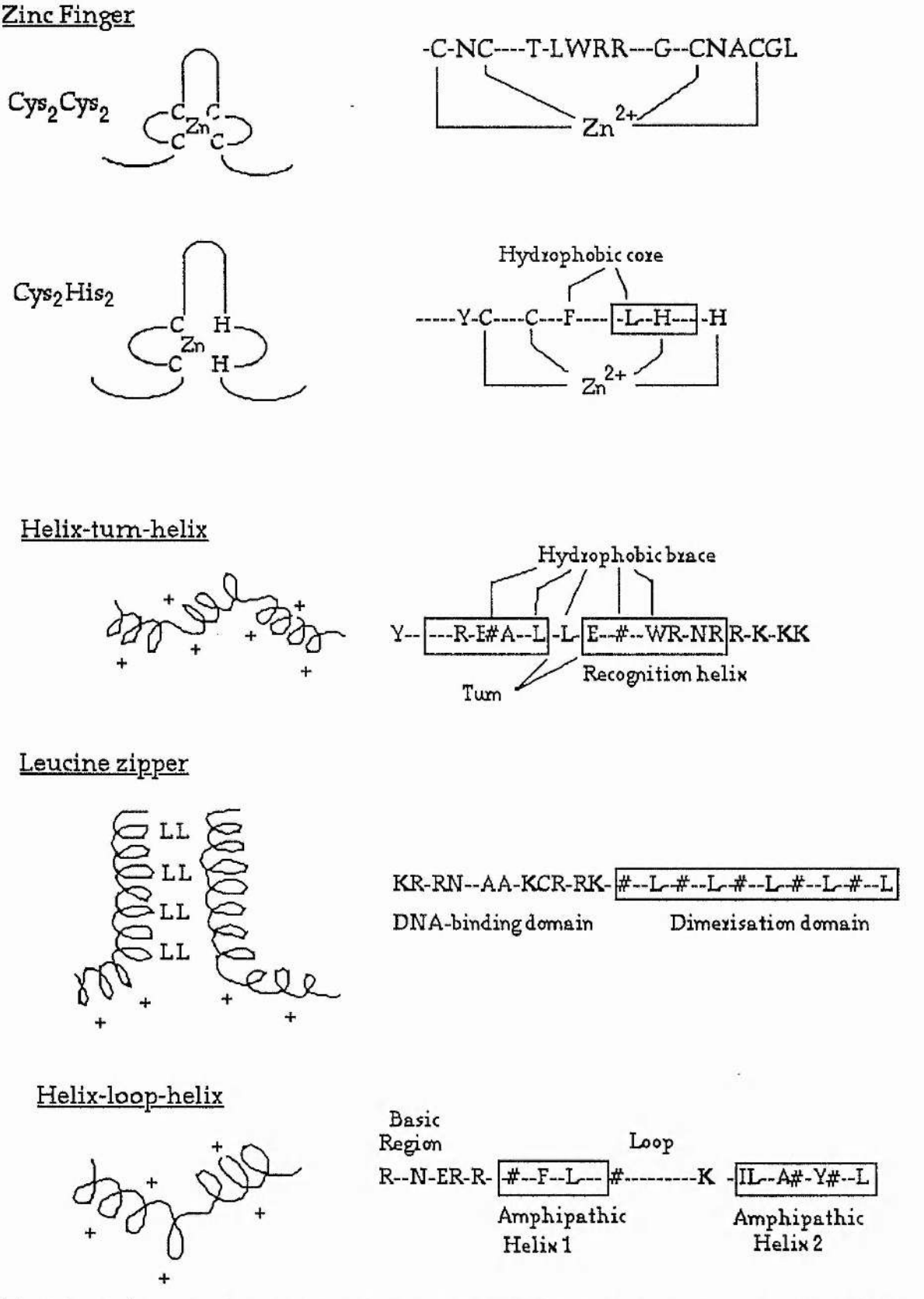


Fig 1.3 Schematic diagrams and structures of the major sequence-specific DNA-binding motifs. Conserved residues of each motif are shown, with the dollar sign representing any hydrophobic AA. Known or presumed α-helices are boxed (from Glover, 1989, Kudla *et al*, 1990 and Mitchell and Tjian, 1989)

Either the zinc atom can be liganded between two cysteine and two histidine residues (Miller *et al*, 1985) or between four cysteine residues (for review see Ptashne, 1988). The zinc finger of the *areA* gene product involved with nitrogen metabolite repression in *A.nidulans* has four cysteine residues (Kudla *et al*, 1990).

Leucine zippers (Fig 1.3) consist of four or five leucine residues, placed seven residues apart, each forming part of an α-helix. The high density of oppositely charged residues within this area results in very stable helical domains. Hydrophobic bonds hold the two helices in an anti-parallel arrangement, hence the name "zipper". The DNA binding site is a highly basic region located N-terminal to the leucine residues (Landschulz *et al*, 1989).

The study of leucine zipper proteins has lead to the realisation that protein transcription factors are often comprised of homo- or heterodimers having different binding or activation potentials (Landschulz *et al*, 1989). The products of the *jun* and *fos* gene families, forming the enhancer binding factor AP-1, can only bind to DNA as dimers. Mutagenesis of these proteins has revealed that the basic region of the leucine zipper is necessary for DNA binding, whereas the zipper itself mediates dimerisation (Turner and Tjian, 1989). The particular nature of the zipper specifies which partnerships are permissible, resulting in repression or activation. For example, a Jun homodimer binds to DNA with low affinity. In contrast, a Jun - Fos heterodimer binds to the same DNA site with high affinity (Jones, 1990).

The helix-loop-helix (Fig 1.3) consists of two helices separated by a loop within the protein structure. The helices are necessary for protein dimerisation whereas the preceding basic region contacts DNA, like the leucine zipper. Although not many examples have been discovered to date, they are present in the Myc family of oncogene proteins (Murre *et al*,

helix-loop-helix motifs and are capable of dimerisation (Jones, 1990). The two activator proteins have an adjacent basic domain permitting DNA binding. In contrast, the two repressor proteins are unable to interact directly with the DNA due to the absence of a basic region. Therefore transcriptional repression occurs when the repressor proteins dimerise with the activator proteins which remain bound to the DNA.

Areas of the C-terminus of regulatory proteins rich in acidic residues are termed "acid blobs". They are conformationally ill-defined with a function irrespective of sequence, provided there are sufficient acidic residues. It is thought acid blobs may stimulate the formation and activation of transcriptional pre-initiation complexes by specifically interacting with the C-terminus of RNA pol II. Within this region a repeating sequence, of TyrSerProThrSerProSer, is capable of offering hydroxyl groups in the formation of hydrogen bonds with the available carboxylates of acidic residues within "acid blobs" (Sigler, 1988). Recently, regions of regulatory proteins rich in proline or glutamine residues have also been demonstrated to activate transciption (Mitchell and Tjian, 1989).

One system of RNA pol II transciptional regulation that has been particularly well studied is that of galactose metabolism in *S.cerevisiae*, for which there are five structural genes. The presence of galactose results in the induction of the genes necessary for galactose metabolism. Conversely, glucose results in carbon catabolite repression. The simplest explanation of how regulation of gene expression occurs is that the *GAL4* gene product, an 881 AA protein, binds to DNA and brings about transcriptional initiation in the presence of galactose and the absence of glucose. However, it was found that only the first 73 AAs were necessary for DNA binding, whereas the C-terminus is involved in transcriptional initiation (Keegan *et al*, 1986). Within the N-terminal 73 AAs of GAL4 is a zinc binuclear cluster specified by six cysteine residues, supporting the

view that this region of GAL4 binds to the DNA. Subsequently, a factor known as GAL80, was shown to bind to GAL4 when there was no inducer, galactose, present. *S.cerevisae* strains either lacking GAL80 or the 30 C-terminal AAs of GAL4 were capable of constitutive activity of the enzymes concerned with galactose metabolism. (Ma and Ptashne, 1987). This led to the hypothesis that GAL80 binds to the C-terminal 30 AAs of GAL4, the latter remaining bound to the regulatory sites of the structural genes. In the presence of the inducer GAL80 dissociates from GAL4 and transciptional factors take its place, resulting in transcriptional initiation (Johnston *et al*, 1987).

1. 3. 2. hnRNA processing

To increase the stability of the hnRNA, and subsequently the mRNA, a structure known as the cap is added to the 5' end of nascent hnRNA chains (Furuichi *et al*, 1977). It is possible that the cap may play a role in translation (Rhoads, 1985).

Further stability of the hnRNA chain is acquired from the poly(A) tail. The efficiency of poly(A) addition could therefore be a determinant in mRNA turnover. Various poly(A) sites have been identified, none as efficient as the more usual, AAUAAA (Wickens, 1990). The absence of convincing GT or T rich sequences downstream of the poly(A) site may lead to lower steady-state mRNA levels (Gil and Proudfoot, 1987).

Some genes make use of alternative poly(A) sites, increasing the diversity of proteins synthesized. The gene encoding human calcitonin has two poly(A) sites. In the thyroid, the poly(A) site at the end of exon 4 is utilised, whereas in brain tissue exon 4 is spliced out resulting in the extension of the transcript to the poly(A) site in exon 6 (Amara *et al* , 1982).

extension of the transcript to the poly(A) site in exon 6 (Amara *et al* , 1982).

hnRNA is found in association with a defined set of abundant proteins, known as hnRNP. Assembly occurs on nascent transcripts leading to the suggestion that hnRNP is the substrate for splicing and 3' end formation (Pederson, 1983).

Splicing of the exons, a further rate-limiting step in mRNA synthesis, usually occurs after the poly(A) tail has been added (Darnell, 1982). Joining different combinations of exons is an important way of increasing the diversity of proteins encoded by one gene (Leff *et al*, 1986 and Breitbart *et al*, 1987). This process, known as alternative splicing, regulates the expression of different gene products under changing circumstances (Smith *et al*, 1989).

1. 3. 3. Transport of mRNA

In eukaryotes mRNA is required to be transported across the nuclear membrane to enable translation to occur in the cytoplasm. Very little is known about nuclear mRNA transport but it has been suggested that it may be differential (Darnell, 1982). hnRNA molecules not fully processed to mRNA will not be transported.

The proteins associated with the mature mRNA in the cytoplasm are different from those of the hnRNP. Therefore the possibility remains that hnRNP organisation may be involved in the transport of mRNA (Dreyfruss, 1986).

Experiments with radiolabelled mRNAs from HeLa cells have demonstrated half-lives as great as 7 and 24 h, the turnover reflecting the replication rate (Singer and Penman, 1973). Likewise, prokaryotic mRNA half-lives are recorded as 2-3 min. The rate of degradation of mRNA in the cytoplasm is important not only in helping to determine its final

steady state level but also in controlling the speed with which changes in the transcription rate are reflected in changes of its cytoplasmic level. mRNA molecules exhibiting a short half-life respond rapidly to changes in gene transcription rates. The poly(A) tail is subject to gradual shortening in the cytoplasm and mRNA degradation may occur after cleavage near the 3' end of the mRNA, resulting in deadenylation (Shapiro *et al*, 1988).

1. 3. 4. Translation

Translation is the process whereby a mature mRNA chain in the cytoplasm is used as a template for the synthesis of a polypeptide chain. The cap acts as a recognition site for a complex of factors including the ribosome (Grifo *et al*, 1983). Before the initiation codon can be isolated by the ribosome complex the mRNA secondary structure must be broken (Kozak, 1983). AAs are carried to the ribosome by tRNAs, the availability of which are thought to be well matched to codon usage and therefore possibly not a major limitation in translation (Hershey *et al*, 1986).

Termination of protein synthesis is promoted by a single release factor recognising all three termination codons, UAA, UAG and UGA. Some proteins may be terminated prematurely, for example, translation of secretory proteins can be blocked at an early stage by the "signal recognition particle" if the appropriate ER components are not present (Walter *et al*, 1984).

1. 3. 5. Post-translation

Post-translational modifications of proteins may occur, depending on the role of the final product. Nearly all proteins synthesized by ribosomes attached to the ER acquire covalently-linked carbohydrates

Insulin has two polypeptide chains that are linked by disulphide bridges but it does not display this structure at the moment of synthesis. Instead the precursor is comprised of a single large polypeptide, proinsulin, which is subsequently cross-linked by intramolecular disulphide bridges and later modified by the removal of peptide segments (Steiner and Oyer, 1967).

A further example is the cleavage of pepsinogen to yield pepsin, the principle proteolytic enzyme in the gastric juice of vertebrates (van Vunakis and Herriot, 1956).

It is evident then, that gene expression is very tightly regulated by a great number of factors creating a very complex system. It is possible to crudely classify the level of regulation at either mRNA accumulation or translation by studies of Northern and Western blots, respectively.

1. 4. 0. NITRATE ASSIMILATION IN *A.NIDULANS*

1. 4. 1. Background

Most members of the bacteria, yeast, filamentous fungi and higher plants have the ability to assimilate nitrate. Much attention has been focussed on the biochemistry, physiology and genetics of this process. The importance of the system is obvious since most of the nitrogen aquisition of these organisms is via the two step reduction of nitrate to ammonium, amounting to 10^4 megatons of nitrate assimilated annually. In addition to nitrate occurring naturally in the environment, nitrate fertilizers are added to crops in order to increase their productivity. There are problems associated with the widespread use of these fertilizers. First, they are produced with fossil fuels and hence have a high cost. Second, nitrates leach into the soil and accumulate in surface and ground waters resulting in algal blooms and more importantly perhaps, the reduction of nitrate to

that of sulphate uptake by *N.crassa* in which there are two permeases, one active in conidia and the other active in mycelia (Marzluf, 1970).

Although Brownlee and Arst (1983) suggested that a functional nitrate reductase enzyme is essential for nitrate uptake, Larsson and Ingermarsson (1989) considered that, in plants, functional nitrate reductase is not a pre-requisite and that free nitrate is the inducer of nitrate uptake.

The *crnA*1 mutation results in a strain that is able to utilise nitrate and is also chlorate resistant (Cove, 1976 b). Chlorate resistance is considered to result from an inability by the strain to synthesize chlorite due to the loss of either chlorate uptake or its reduction to chlorite.

1. 4. 3. Nitrate reductase

Synthesis of the enzyme, NR is induced by nitrate and repressed by ammonium (Cove, 1966) or the lack of glucose (Hynes, 1973). There are three closely related forms of nitrate reductase: NADH NR (E.C. 1.6.6.1.) is the most common form in higher plants and algae, some of which also contain NAD(P)H NR (E.C. 1.6.6.2.), whilst NADPH NR (E.C. 1.6.6.3.) is found in fungi. NR is a soluble, multi-centre redox enzyme catalysing the two electron reduction of nitrate to nitrite using pyridine nucleotides as the electron donor. It is a homodimer with each subunit composed of a polypeptide with a molecular weight of 100 kDa and three co-factors in a 1:1:1 ratio (Redinbaugh and Campbell, 1985). These co-factors are FAD, heme Fe and molybdenum (Mo)-pterin which have binding sites in a linear arrangement along the length of the 900 AA polypeptide (Fig 1.4) (Campbell and Kinghorn, 1989). AA sequence alignments between *A.nidulans*, bacteria and plant NRs have demonstrated that the *A.nidulans* NR is more closely related to the form generally present in

toxic compounds, by bacteria in the human gut, possibly leading to carcinomas. Finally, there is the problem of denitrification of nitrate in the soil to nitrous oxide, a greenhouse gas (Campbell and Kinghorn, 1990).

Current aspects of nitrate assimilation research are based upon extending our basic understanding of the pathway and hopefully enhancing the efficiency of the process to improve crop yield, thereby solving some of the economic, ecological and medical problems associated with nitrate fertilisation.

The nitrate assimilatory pathway of *A.nidulans* is particularly well characterised at the biochemical and genetic level (Cove, 1979). Although it is not important commercially, the system in *A.nidulans* may hopefully act as an applicable model for that in higher plants. It may even be possible for the *A.nidulans* genes to function in plants although this seems unlikely due to the differences in gene structure of higher eukaryotes. Despite this, nitrate assimilation in *A.nidulans* is of direct interest as a means of expanding our understanding of gene regulation.

1. 4. 2. Nitrate permease

It has been demonstrated that net nitrate uptake in *A.nidulans* is inhibited by uncouplers indicating that a proton gradient across the plasma membrane is required for uptake (Brownlee and Arst, 1983). Strains carrying a mutation in the gene encoding the nitrate permease, *crnA*, are able to assimilate nitrate suggesting the existence of more than one nitrate uptake system. The broad pH dependency and the kinetics of nitrate uptake suggest systems with different affinities for nitrate. Brownlee and Arst (1983) observed that the *crnA*1 mutation reduced net nitrate uptake in conidia and young mycelia but not in older mycelia. In this respect, the *A.nidulans* uptake system may bear some resemblance to

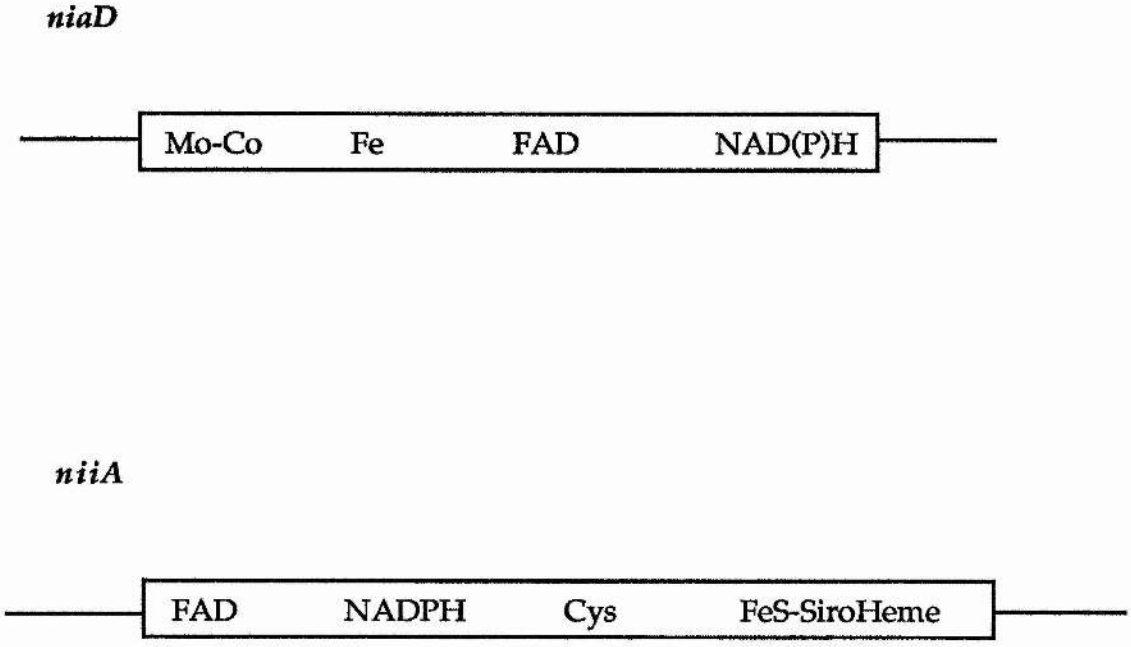


Figure 1.4 A schematic diagram of the *niaD* and *niiA* encoded polypeptides, indicating the relative position of the co-factor-containing functional domains (Campbell and Kinghorn, 1990).

plants (Kinghorn and Campbell , 1989).

Several mutations within the *niaD* gene, encoding the NR apoenzyme, result in constitutive nitrite reductase and hydroxylamine reductase activities (MacDonald and Cove, 1974) and also chlorate resistance (Cove, 1976 a). In addition, there is a group of *niaD* mutant strains posessing thermolabile nitrate reductase, confirming that *niaD* is indeed the structural gene for NR (MacDonald and Cove, 1974).

1. 4. 4. The *cnx* genes

Strains with mutations within any of the *cnx* genes are unable to utilise nitrate or hypoxanthine but they are able to utilise nitrite (Cove and Pateman, 1963) and can be selected on the basis of chlorate resistance (Cove, 1976 a). There are eight *cnx* loci; *cnxA, B* and *C* are tightly linked on chromosome VIII, with *cnxE - H* (Cove, 1979) and *J* (Arst *et al*, 1982) arranged throughout the genome. Pateman *et al* (1964) suggested they may determine a common component necessary for NR and xanthine dehydrogenase. Molybdate supplied to cultures of *cnxE* mutants partially restored all three enzyme activities (Arst *et al*, 1970). This augments the theory that the *cnxE* gene may be responsible for the insertion of molybdenum into a co-factor encoded by the other *cnx* genes. The function of the other *cnx* genes remains uncertain.

1. 4. 5. Nitrite reductase

It has been shown that the enzyme, NiR is inducible by nitrate and nitrite but is repressed by ammonium (Cove, 1966). There are two forms of NiR. Higher plant NiR (E.C. 1.7.7.1.) utilises reduced ferredoxin as the electron donor. In contrast, the electron donor of bacterial and fungal NiR (E.C.1.6.6.4.) catalysing the six electron reduction from nitrite to ammonium, is NAD(P)H. Like NR, NiR is soluble and contains siroheme Fe, FeS and FAD as redox centres laid out in a linear arrangement (Fig 1.4). The *N.crassa* enzyme is a homodimer with a polypeptide of molecular weight 290 kDa (Lafferty and Garrett, 1974), and is also capable of NADPH activity using ferricyanide and cytochrome c as electron acceptors.

In contrast to NR, AA alignments reveal that fungal NiR has greater similarity with bacterial NiR than plant NiR (Kinghorn and Campbell, 1989).

Strains carrying mutations mapping at the *niiA* locus, the gene encoding NiR, are unable to utilise either nitrate or nitrite as a nitrogen source (Pateman *et al*, 1967). Such strains have identical growth characteristics to the *nirA*1 strain (1.5.7.a). The major distinguishing feature is that the former excrete nitrite and it is possible to stain for this (Cove, 1976 a). In addition *nirA*$^-$ strains are chlorate sensitive irrespective of the nitrogen source. Rand and Arst (1977) showed that NiR was less stable in a temperature sensitive *niiA* mutant strain, which is consistent with the *niiA* gene being the structural gene for NiR.

Mutations at the *nihA* locus on linkage group I result in nitrite toxicity. It was shown that this was not due to a defect in the NiR enzyme and the possibility was raised that the mutation leads to an increase in efficiency of nitrite uptake which is not suppressed by ammonium (Pombeiro *et al*, 1983).

1. 4. 6. Structure of the nitrate gene cluster

Several genes of related function are arranged contiguously in filamentous fungi, for example the genes for quinic acid utilisation in *N.crassa* and *A.nidulans*, *qa* (Giles *et al*, 1985) and *qut* (Da Silva *et al*, 1986) respectively, the *alcA* and *alcR* genes for ethanol utilisation in *A.nidulans* (Pateman *et al*, 1983) and the *crnA*, *niiA* and *niaD* genes for nitrate utilisation in *A.nidulans* (Tomsett and Cove, 1979 and Johnstone *et al*, 1990). Each gene within a cluster is transcribed separately differing from the prokaryotic operon in which a single polycistronic transcript is synthesized, such as the genes of the lactose operon of *E. coli* (Jacob and Monod, 1961). That pairs of genes within clusters are often divergently transcribed perhaps suggests a purpose for this type of gene arrangement.

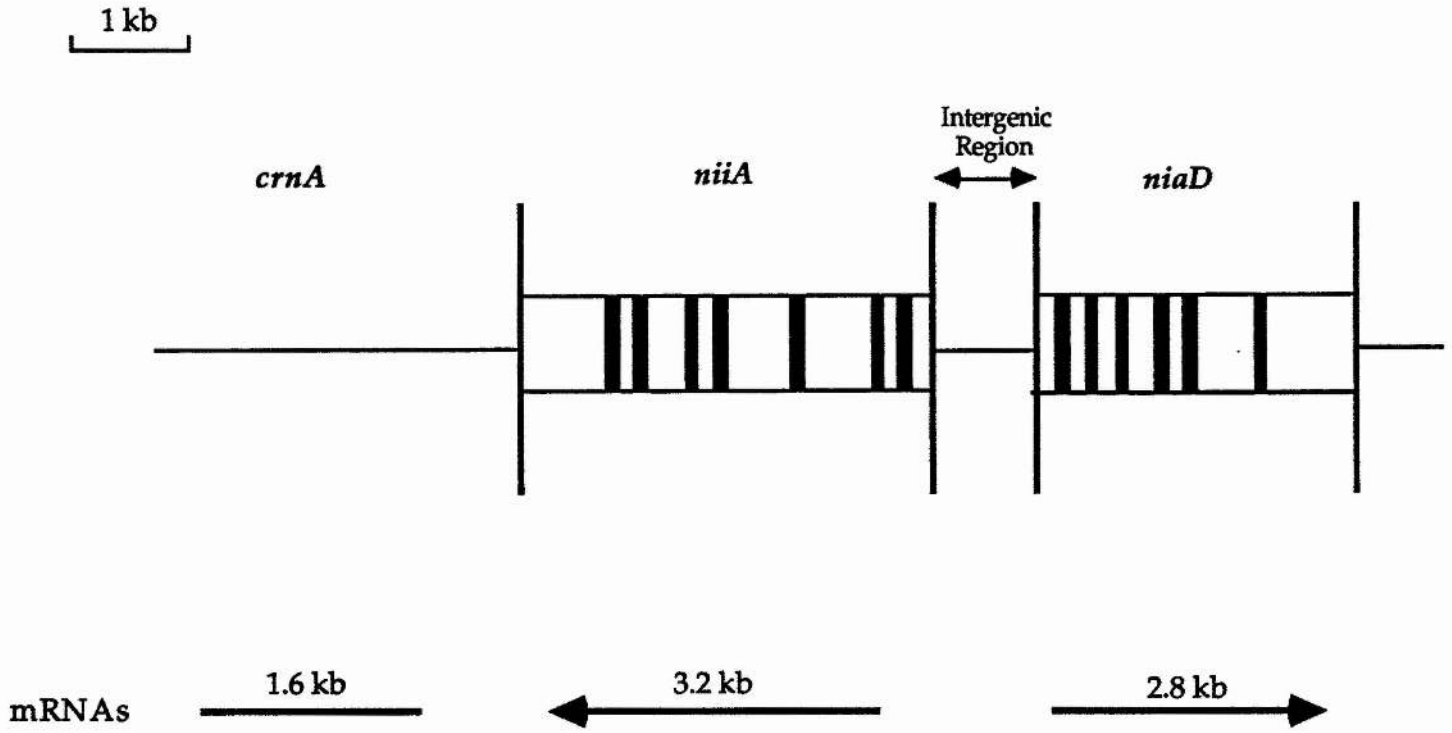


Figure 1.5 Schematic diagram of the *Aspergillus nidulans* nitrate gene cluster. The extent of the *niiA* and *niaD* genes is represented with a vertical line. The white boxes represent the exons and the position of introns is indicated by the black boxes. The size of the transcripts and the direction of transcription for each gene is indicated by the horizontal arrows. Information regarding the *crnA* gene is limited at this stage.

It is thought that upstream elements may be able to function bi-directionally reducing the number of elements required for gene activation. In addition, this results in aggregating the activator binding sites, increasing the effective concentration of activator which may otherwise be limiting. The origin of gene clusters may be a result of divergent evolution from the prokaryotic operon (Geever *et al*, 1989). However this raises the question as to why, for example, the nitrate assimilatory genes of *N.crassa* are not arranged in a cluster when those of its close relative, *A.nidulans* are.

The three genes *crnA*, *niaD* and *niiA* on chromosome VIII were isolated on an 8 kb stretch of DNA that was found to complement an *A.nidulans* strain lacking NiR activity (Johnstone *et al*, 1985). The nitrate cluster is situated on chromosome VIII in the arrangement of *crnA-niiA-niaD* (Fig 1.5). There is less than 1 kb of DNA between each gene, leading to the consideration that, despite there being no eukaryotic precedents, they may form an operon. However, a mutation, *nis-5*, involving a translocation from chromosome II is tightly linked to the *niiA* gene and was determined to only affect the expression of the *niiA* gene providing initial evidence that the *niiA* and *niaD* genes are transcribed separately (Rand and Arst, 1977 and Arst *et al*, 1979). Further to this, a separate transcript has now been identified for each gene, with sizes of 1.6 kb, 3.2 kb and 2.8 kb for *crnA*, *niiA* and *niaD* respectively (Fig 1.5) (Johnstone *et al*, 1990).

The *niiA* and *niaD* genes have been shown to be divergently transcribed (Johnstone *et al*, 1990) in common with the members of other gene clusters. The nucleotide sequences of both, including the intergenic region, have been determined and their structural details evaluated (1.2.1). The sizes of the two genes are 3.2 kb and 3.7 kb for *niaD* and *niiA* respectively, the latter having seven introns, one more than *niaD* (Fig

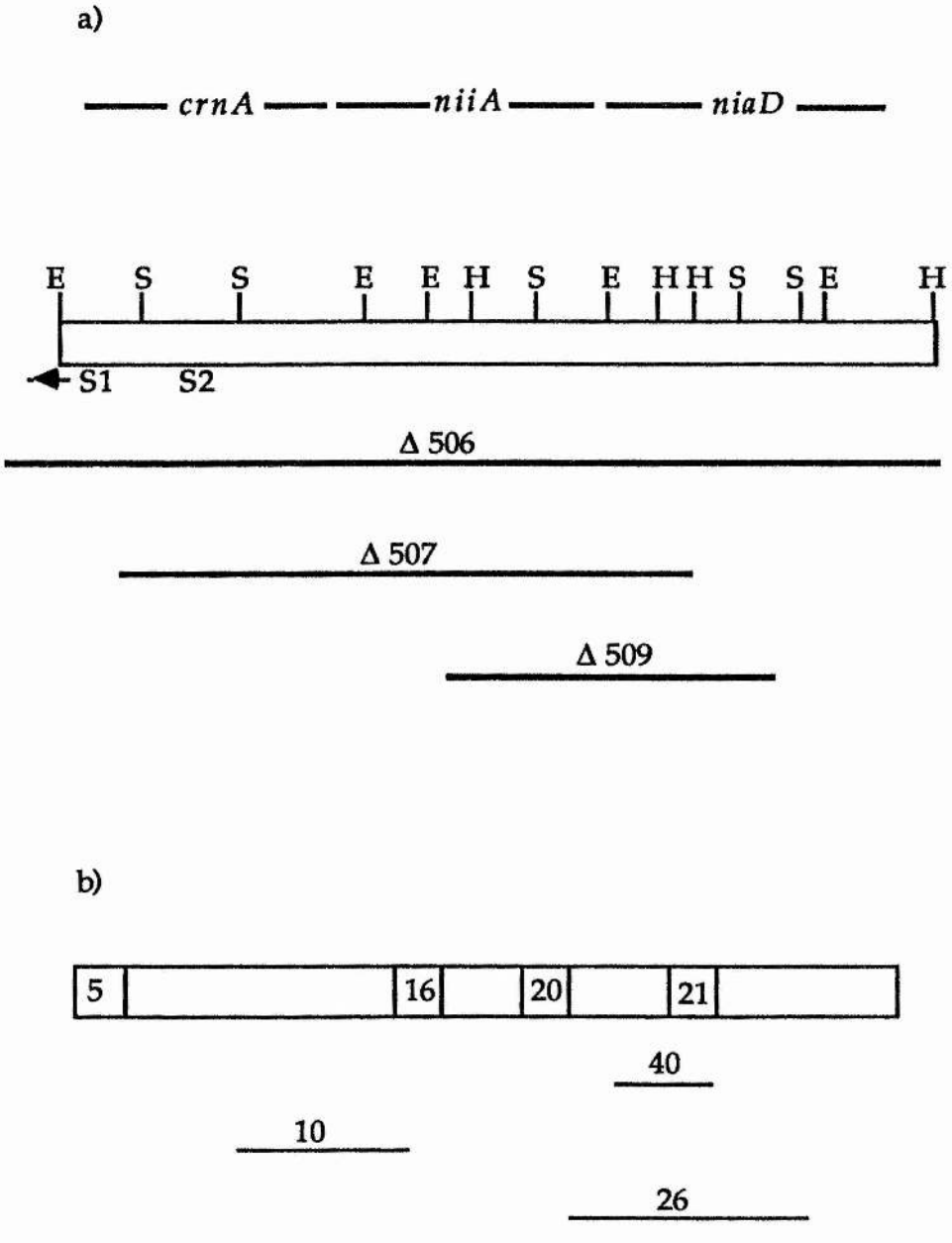


Figure 1.6 a) Restriction map of the *crnA-niiA-niaD* gene cluster. Horizontal lines represent the extent of each deletion. The restriction sites shown are E; *Eco*RI, S; *Sal*I and H; *Hind*III. S1 and S2 are the *Sal*I fragments within the *crnA* gene. b) Schematic diagram of the *niaD* gene indicating the position of the small deletions, represented with horizontal lines. The approximate position of a few point mutations are shown. The diagrams are adapted from Tomsett and Cove (1979) and Johnstone *et al* (1990).

1.5). Several mutations within the cluster have been characterised (Tomsett and Cove, 1979) (Fig 1.6) permitting the effect of these mutations on the expression of each, individual gene to be determined. In addition, it should be possible to locate the functional domains of the *niaD* gene for which there are a greater number of mutant strains available

1. 4. 7. Regulation of nitrate assimilation

There are at least two positive regulatory genes involved with nitrate assimilation in *A.nidulans.*

The *nirA* gene is unlinked to the nitrate cluster genes on chromosome VIII (Clutterbuck, 1984). The mode of action of its product is pathway specific and necessary for the induction of the structural genes, in the presence of nitrate and absence of ammonium. However, it is not thought to play a part in the regulation of the *crnA* gene (Brownlee and Arst, 1983).

The *areA* gene, is situated on chromosome III (Clutterbuck, 1984). Its product has a wider regulatory role and in the presence of ammonium, is required for nitrogen metabolite repression of genes involved with the assimilation of a variety of nitrogen sources.

The two regulatory proteins, products of the *nirA* and *areA* genes, may act in a similar fashion to that of the *GAL4* gene involved with the transcriptional activation of galactose metabolism in *S.cerevisiae* (1.3.1). One domain of the protein may be responsible for DNA binding with another involved in transcriptional activation (Scazzocchio and Arst, 1989). In particular, the C-terminus of the *areA* gene product may interact with glutamine. This effector, the end product of nitrogen metabolism, inactivates the protein and prevents further transcription of the structural genes. It would appear therefore, that the DNA-binding

domain of this protein is central and N-terminal to the region responsible for transcriptional activation (1.3.1.) (Kudla *et al*, 1990).

Three possible mechanisms for the way in which the *nirA* and *areA* gene products may interact with each other and the intergenic region have been suggested (Scazzocchio and Arst, 1989). First, the product of the *areA* gene ligands to the *nirA* gene product and this complex binds to *cis*-acting elements in the intergenic region. The second model considers that the two gene products may bind to individual *cis*-elements without interacting directly with each other. Finally, the *areA* gene product may be necessary for transcription of the *nirA* gene, in which case only the *nirA* gene product binds to the *cis*-acting regions of the *niaD* and *niiA* genes. These mechanisms may not be mutually exclusive.

It is thought that the NR holoenzyme has an autoregulatory role (Cove and Pateman, 1969). A model of how NR may interact with the *nirA* gene product and nitrate to exert an effect on regulation is provided by Cove (1979), (Fig 1.7). Transcription of the structural genes occurs when nitrate is present and complexed with NR, allowing the *nirA* gene product to interact freely with the *cis*-acting sites of the intergenic region. When nitrate is absent, NR prevents the *nirA* gene product from interacting with the *cis*-elements either by direct competition or by complexing with it. It is thought that the rare $nirA^{c}$ mutation (Pateman and Cove, 1967), resulting in constitutive *niiA* and *niaD* gene expression, in the absence of the inducer, is a result of the disruption of the NR binding domain of the *nirA* gene product (Scazzocchio and Arst, 1989).

1. 4. 7. a. The *nirA* gene

A mutation in the *nirA* gene can lead to the inability to utilise nitrate or nitrite as a nitrogen source. Such mutations designated $nirA^{-}$

have the same phenotype as *niiA* mutants but can be distinguished, as previously explained (1.4.5.) (Cove, 1976 b). The rare mutant allele, *nirA*^{c}, resulting in constitutive NR activity (Pateman and Cove, 1967), maps at the *nirA* allele (Cove, 1970) indicating that the gene has a regulatory role in nitrate assimilation. Since most mutations lead to non-inducibility of the structural genes and *nirA*^{c} mutations are rare, it was concluded that the *nirA* gene product has a positive regulatory role (Cove, 1979).

NR activity in diploids was found to be greatest in *nirA*^{c}/*nirA*^{c} strains but similar levels were found in *nirA*^{c}/*nirA*$^{+}$ and *nirA*^{c}/*nirA*$^{-}$ strains, indicating that the *nirA* gene product is present in near limiting concentrations (Cove, 1979). Another rare mutant, *nirA*^{d} (Rand and Arst, 1978), later named *nirA*$^{c/d}$ (Tollervey and Arst, 1981), was found to result in constitutivity and to partially alleviate the need for a functional *areA* gene product. It was thus postulated that the *nirA* gene consists of at least two genetically separable domains, one in which mutations lead to constitutivity, possibly encoding a co-inducer (NR) binding region (Scazzocchio and Arst, 1989) and another in which mutations lead to nitrogen metabolite derepression, possibly encoding a region interacting with the *areA* gene product (Tollervey and Arst, 1978).

1. 4. 7. b. The *areA* gene

Several different phenotypes are associated with various mutant alleles of the *areA* gene, although only two general classes are recognised. *areA*^{r} mutations result in the inability to utilise nitrogen sources that are normally repressible by ammonium and *areA*^{d} mutations lead to derepression of either one or more activity that is normally repressed by ammonium (Arst and Cove, 1973). It is possible that repression mechanisms have a wider regulatory domain in order to prevent utilisation of less favoured nutrients, when preferred nutrients are

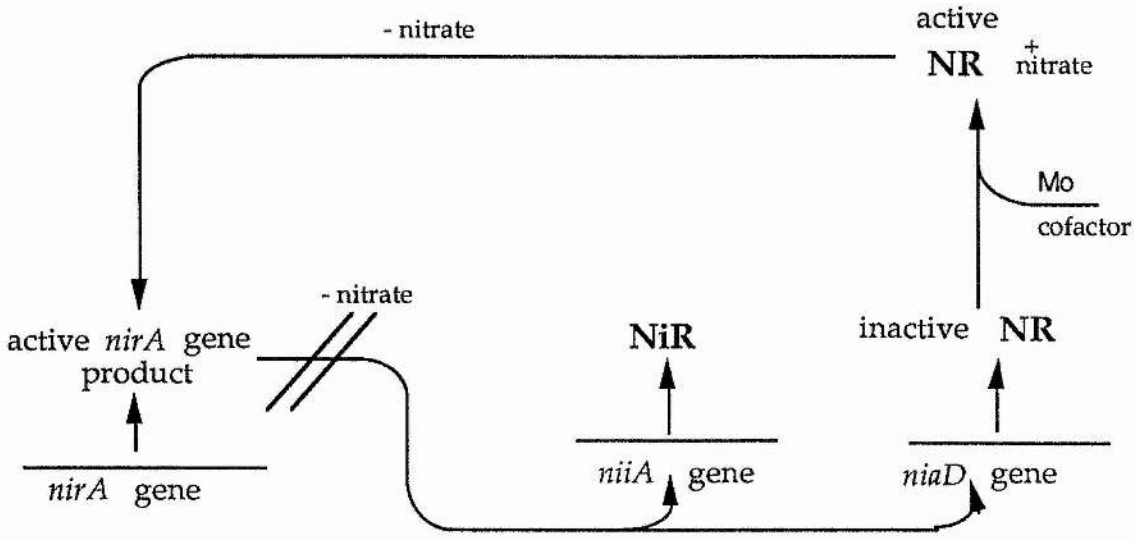


Figure 1.7 The regulation of *niiA* and *niaD* gene expression by the *nirA* gene product, demonstrating the proposed interaction with the NR enzyme.

---

available. This would involve the expression of fewer enzymes and permeases (Arst, 1984).

Due to the range and the consequence of *areA* mutations it was suggested that the *areA* gene product is required for the expression of genes that are normally subject to ammonium repression and that the presence of ammonium would inactivate the *areA* gene product (Arst and Cove, 1973). However, it is probable that glutamine, the main end product of ammonium assimilation, is the repressor (Scazzocchio and Arst, 1989).

The *areA*-associated *xprD*1 mutation (Cohen, 1972) truncates the 124 C-terminal AA residues of the *areA* encoded polypeptide (Kudla *et al*, 1990) resulting in nitrogen metabolite derepression. It was shown that neither *areA*$^-$/*nirA*^{c} nor *xprD*1/*nirA*$^-$ double mutants were able to grow on nitrate indicating that both control genes need to be functional in order to obtain expression of the *niaD* and *niiA* genes. Consequently, it was concluded that nitrate induction and ammonium repression of NR and NiR are independent (Arst and Cove, 1973).

The mutation *areA*19 results in loss of *areA* gene function. It is possible to restore this by introducing the regulatory gene *nit-2*, from *N.crassa*, by a gene-mediated transformation system (for review see Fincham, 1989) (2.2.2). These results imply that although the two genes *areA* and *nit-2* are from different organisms and bear no relationship at the nucleotide level, they have a similar regulatory role, indicating that their functional motifs may have been conserved in evolution (Davis and Hynes, 1987). Despite this, it was observed that the *nit-2* transformants had derepressed NR activity, suggesting that the product of a second *N.crassa* regulatory gene, *nmr* (Sorger *et al*, 1989), is required for total nitrogen metabolite repression. This raises the question of whether other genes involved with nitrate assimilation are interchangeable between the two organisms.

Recently, the *areA* gene has been cloned (Caddick *et al*, 1986) and sequenced (Kudla *et al*, 1990). The protein consists of 719 AAs of which only residues 343-595 are essential to gene activation function. Several features common to regulatory proteins were found on the *areA* gene product including a zinc-finger and "acid blobs" (Sigler, 1988), responsible for DNA binding and transcriptional activation respectively (1.3.1).

1. 5. 0. TRANSPORT SYSTEMS

The information provided above clearly shows an imbalance in our knowledge of nitrate assimilation and that nitrate transport studies in *A.nidulans* require further attention. The size of the gene involved, how it is regulated and the way in which the permease functions are all questions that have to be addressed. In relation to this a survey of other transport genes and the permeases they encode should provide a background for the nitrate permease study.

Living cells are separated from their environment by a membrane which confers individuality and acts as a highly selective permeability barrier. Nutrients from the environment have to cross this physical barrier before being utilised by the cell.

A detailed account of the structure of membranes is discussed by Dewey and Barr (1970). Briefly, a membrane consists of a bilayer of lipids, the majority of which are phospholipids. The interior of the membrane is hydrophobic, with the fatty acid side chains of each lipid molecule orientated inwards and the polar phosphate side chains at the surface of the membrane. Specific proteins found intercalated within the membrane structure can be classified as monotopic, bitopic or polytopic according to the number of membrane associated domains (1.5.2.) (Blobel, 1980).

More than thirty years ago Rickenberg *et al* (1956) recognised that proteins are involved in the accumulation of solutes inside cells, suggesting that transport of certain solutes is an active process and not just a case of passive diffusion across membranes. The *E.coli* lactose transporter (Kaback, 1983) was the first permease to be recognised and since then over forty different transport systems have been recognised in *E.coli* alone whilst several in other organisms have also been identified.

1. 5. 1. Mechanisms of transport

In bacteria, transport systems can be classified into different groups depending on their mechanism of action (Hengge and Boos, 1983). Unfortunately, there is little evidence to suggest that any of these categories will be recognised in eukaryotes, with the possible exception of transport systems coupled to ion gradients.

The phosphotransferase system is responsible for the unidirectional transport of various sugars and sugar-alcohols in *E.coli* and *Salmonella typhimurium* (Postma and Roseman, 1976). It is coupled obligatorily to phosphorylation of the substrate with phospho-enolpyruvate as the phosphoryl donor, thereby changing the substrate as it is transported; an unusual feature of transport systems.

Binding proteins are located in the periplasm between the outer and inner membrane of Gram-negative bacteria. It is thought that these proteins are the major recognition sites and carriers of multi-component transport systems (Ames and Higgins, 1983), such as histidine transport in *S.typhimurium* (Ames, 1985) and maltose transport of *E.coli* (Hengge and Boos, 1983). Other protein components necessary for binding-protein mediated transport are located in the inner membrane. The periplasmic protein has a recognition site for these and also for the substrate (Fig 1.8). Features common to this type of transport system include a high affinity with concentration gradients of up to $1 : 10^4$ (out : in), an ATP-dependent unidirectional pump (1.5.2.b) and no modification of the substrate during translocation.

The accumulation of solutes by many transport systems is not linked directly to energy consumption. Instead, energy is acquired by exploiting a pre-existing ion gradient, such as sodium or hydrogen, in order to drive solute transport against a concentration gradient. An example of such a system is lactose transport in *E.coli* (Kaback, 1983). The

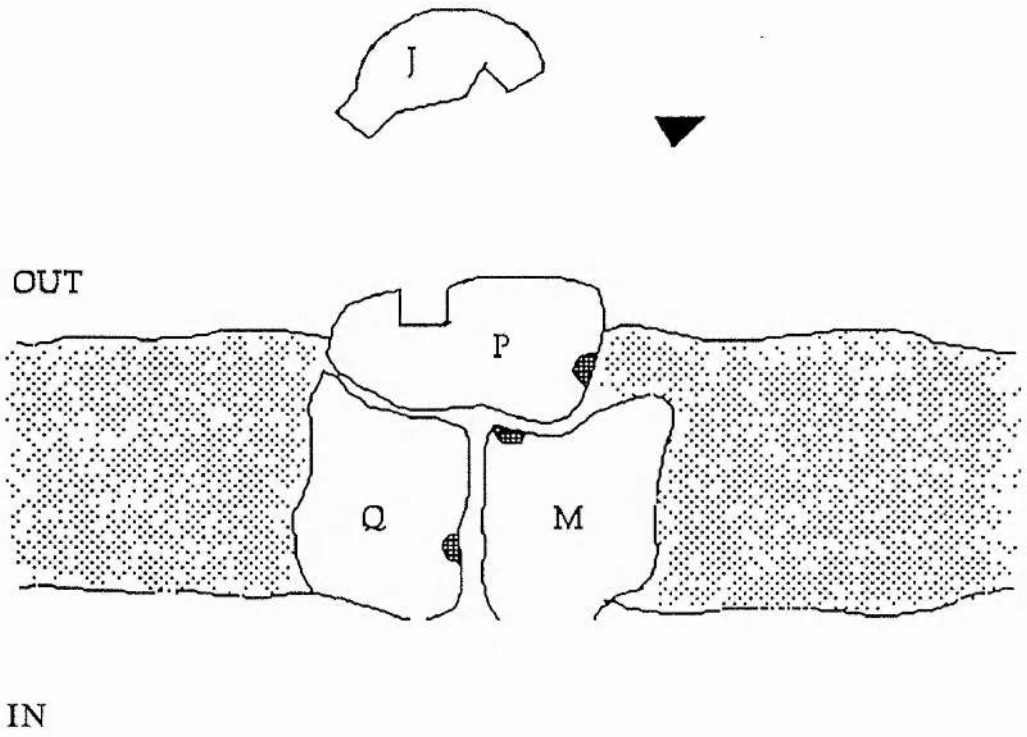


Figure 1.8 Diagramatic representation of the histidine transport system in *Salmonella typhimurium*. The *hisJ* gene product is a periplasmic transport protein, binding with histidine and carrying it across the membrane via a mechanism involving the *hisP*, *hisQ* and *hisM* gene products (Ames, 1985).

"lac" permease utilises the energy released by the translocation of an H^+ ion across the membrane in order to drive the transport of lactose, resulting in accumulation within the cell.

The specific diffusion pore is responsible for the translocation of a compound across a membrane by a protein-mediated diffusion process. Only one such system is known, that of glycerol transport in *E.coli* (Sanno *et al*, 1968). The protein known as a facilitator shows specificity for glycerol and other non-cyclic polyalcohols.

High affinity outer membrane receptors of *E.coli* recognise iron complexes and vitamin B-12 (Braun and Hantke, 1981). These receptors appear to lead to accumulation of the substrate within the periplasm and may require energy input for dissociation of substrate from receptor protein.

To address the problem of substrate specificity of transport proteins and how a solute crosses a membrane several transport proteins of *E. coli* were examined (Ferenci, 1989). No evidence was found for separate binding sites at the external or internal faces of the proteins or one that extended across the whole depth and neither was there evidence for the rotation of binding from one side to the other. However, Lodish (1988) argues that transport involves the alternation of inward and outward facing solute binding sites, the protein undergoing a conformational change.

1. 5. 2. Features of transport proteins

The sizes of transport proteins range from 37 kDa, for the *E.coli* galactose transporter (Macpherson *et al*, 1983), to 97 kDa, the *S.cerevisiae* glucose transporter (Celenza *et al*, 1988). At the AA level permeases appear to have very little similarity with each other, in general. However, there are exceptions; the glucose transporters of the human hepatoma and the rat brain cells share substantial AA identity with both the arabinose and xylose transporters of *E.coli* and the glucose transporter of *S.cerevisiae* (Celenza *et al*, 1988). Additionally, the *A.nidulans* proline transporter and the arginine and histidine transporters of *S.cerevisiae* all show some sequence similarity (Sophianopoulou and Scazzocchio, 1989).

Despite the general lack of sequence identity there are several features that are common to permeases.

1. 5. 2. a. Hydrophobic membrane-spanning domains

These domains are common to all transport proteins. Membrane spanning regions often form α-helices, consistent with this being the most enèrgetically favourable configuration for the arrangement of AAs in membranes. 21 AAs were thought to be necessary to allow the protein to span the membrane, as this results in an α-helix of sufficient length to cross the lipid bilayer. However, it has since been shown that shorter domains of 18, 16 or 14 AAs can assume a transmembrane configuration (Adams and Rose, 1985). It is thought that these shorter domains are responsible for solute binding or pore formation and are located in the centre of the protein structure. The surrounding transmembrane domains of 21 AAs are necessary for anchoring the protein within the bilayer, creating thermodynamically favourable conditions for the inner, shorter and more unstable domains of the protein.

There are several methods available for calculating which windows of 21 AAs are those likely to span the membrane. A window represents a group of 21 AAs from 1 - 21, 2 - 22, 3 - 23 and so on throughout the length of the protein. Engleman *et al* (1986) assigns a value to each AA depending on the free energy of transfer from water into oil, apolar amino acids having lower values. Alternatively, a value for each AA can be determined depending on its "hydrophobicity" or its affinity for the interior of the membrane (Kyte and Doolittle 1982 and Eisenberg *et al*, 1984) (Table 1.1). These hydrophobic AAs are leucine, isoleucine, valine, phenylalanine, alanine, methionine and tryptophan and make up the majority of each membrane-spanning domain. It is very energetically unfavourable to transfer polar AAs into the hydrophobic environment of membranes and hence they are very rarely found in membrane spanning regions. Theoretically, it is possible to

Table 1.1 The hydrophobicity scales of the amino acids, as determined by Eisenberg *et al* (1984). The normalised consensus values used for all calculations in this study have been scaled to have a mean of 0.00 and a standard deviation of 1.00. The positive index amino acids have the greater hydrophobicity.

Residue	Consensus	Normalised Consensus
Arginine	-1.76	+2.53
Lysine	-1.10	+1.50
Aspartic acid	-0.72	+0.90
Glutamine	-0.69	+0.85
Asparagine	-0.64	+0.78
Glutamic acid	-0.62	+0.74
Histidine	-0.40	+0.40
Serine	-0.26	+0.18
Threonine	-0.18	-0.05
Proline	-0.07	-0.12
Tyrosine	0.02	-0.26
Cysteine	0.04	-0.29
Glycine	0.16	-0.48
Alanine	0.25	-0.62
Methionine	0.26	-0.64
Tryptophan	0.37	-0.81
Leucine	0.53	-1.06
Valine	0.54	-1.08
Phenylalanine	0.61	-1.19
Isoleucine	0.73	-1.38

determine whether each membrane-spanning region is either globular, surface-seeking or membrane buried (Eisenberg *et al*, 1984) by plotting the hydrophobic moment of each helix as a function of its hydrophobicity. The hydrophobic moment measures the tendency of the helix to seek a surface between hydrophilic and hydrophobic phases. Alternatively, the hydrophobicity measures the helix's affinity for the membrane interior. Transport proteins have many hydrophobic domains, the purine-cytosine transporter of *S.cerevisiae* has eight hydrophobic regions (Weber *et al*,

1990) compared with the twelve domains of the phosphate transporter of *N.crassa* (Mann *et al*, 1989).

1. 5. 2. b. ATP-binding sites

Two short AA sequence motifs, identified by Walker *et al* (1982), are found in many nucleotide-binding proteins. It was proposed that those bacterial transport systems with periplasmic binding-proteins are directly coupled to ATP utilisation and may possess these motifs. The AA sequences of transport proteins including those encoded by the genes *malK*, the maltose transporter of *E.coli* (Hengge and Boos, 1983) and *hisP*, the histidine transporter of *S.typhimurium* (Ames, 1985) were aligned against the AA sequences of proteins involved with cell division, DNA repair and multidrug resistance. The ATP-binding motifs were common to all these proteins, showing that they are directly involved with the uncoupling of ATP (Higgins *et al*, 1988). Motif A consists of the consensus GlyXXGlyXGlyLysThr, in which only the lysine residue is postulated to be important in nucleotide binding. In motif B an aspartate residue binds MgATP, water being excluded by the preceding four hydrophobic AAs.

1. 5. 2. c. Membrane-buried proline residues

A survey of membrane-spanning regions of proteins showed that nearly all the transport proteins examined had membrane-buried proline residues. In constrast, non-transport membrane proteins were largely devoid of them, (Brandl and Deber, 1986). The fact that proline residues are found within membrane-spanning domains at all is surprising since insertion of a proline residue into a membrane is energetically unfavourable compared with a non-polar AA. As a result of relatively rigid backbone angles and no amide proton for donation in hydrogen bonds proline would be expected to break the classical α-helical structure

of trans-membrane domains. A peptide bond between proline and an unspecified AA within a membrane results in redirection of the protein chain and hence one purpose of membrane-buried proline residues is proposed to be an involvement in the regulation of a transport channel (Brandl and Deber, 1986). A second function was suggested to be a direct involvement in the transport of H^+ ions in transporters that operate a symport mechanism. The oxygen atom within the carbonyl group of a peptide bond between a proline and a second AA has relatively high basicity with the capability of acting as a liganding site for positively charged species.

1. 5. 2. d. Glycosylation sites

Many transport proteins have been shown to have the motif AsnXSer or AsnXThr the site for glycosylation. The asparagine residue is important for the attachment of the N-oligosaccharide. Two such motifs are present in the purine-cytosine transporter (Weber *et al*, 1990) and four in the allantoate transporter (Rai *et al*, 1988) of *S.cerevisiae*. One is also found in the human glucose transporter (Mueckler *et al*, 1985). It is believed that most proteins sequestered in the lumen of the ER before being secreted or inserted into the membrane are glycoproteins (Lodish, 1988), and require these motifs. The function of the oligosaccharide chain appears to differ, although suggested roles include cell surface receptors, signals for protein targetting and a means of cell to cell interaction (Paulson, 1989).

1. 5. 3. Protein assembly

Various 'topogenic' sequences within the polypeptide of a membrane protein are proposed to serve certain functions for insertion and assembly, creating the architecture of the protein in the membrane

(Blobel, 1980). The first hydrophobic domain, the "signal sequence", is responsible for the translocation of a protein across a membrane. In contrast, the second hydrophobic domain, the "stop-transfer sequence", interrupts translocation, whilst the third re-initiates translocation, and so on.

Evidence has been provided by Wessels and Spiess (1988) that protein insertion into membranes occurs co-translationally. The ribosome becomes attached to a docking particle on the membrane, where protein synthesis ensues. The newly assembled signal sequence is recognised by a signal recognition particle on the membrane and becomes inserted at this site. Hydrolysis of GTP may play a role in enhancing the specificity of this recognition process (Rapoport, 1990).

An opposing theory is that of post-translational insertion (Verner and Schatz, 1988) and in particular two mechanisms have been proposed. The membrane trigger hypothesis, put forward by Wickner (1979) depends upon the ability of the lipid bilayer to trigger the folding of a polypeptide into a suitable conformation such that it may be incorporated into the membrane. It is not clear how this may occur but the signal sequence is postulated to be necessary for activating the protein assembly. Alternatively, the helical hairpin hypothesis (Engelman and Steitz, 1981) depends upon the formation of two α-helices, arranged in a hairpin structure, the signal sequence being the first. This spontaneously partitions from the aqueous cytoplasm into the energetically more favourable environment of the lipid bilayer, with the remainder of the protein being inserted in such a way as determined by subsequent non-polar α-helices. A recent review suggests that the mechanism of insertion, whether co-translational or post-translational, is dependent upon the membrane type (Wickner, 1989).

The orientation of a protein within a membrane is considered to be determined by the position of arginine and lysine residues within hydrophilic regions. Generally, a net positive charge can be found on the cytoplasmic side of a membrane-spanning stretch (von Heijne and Gavel, 1988). In contrast, particularly in *E.coli*, the arrangement of acidic residues is less important (Nilsson and von Heijne, 1990). Further, the overall structure of the protein is proposed to be stabilised by interactions between adjacent hydrophilic and hydrophobic domains (Boyd and Beckwith, 1990).

The above account suggests that there appear to be conflicting ideas about many aspects of transport processes and much work has to be done to extend our understanding.

1. 6. 0. EXPERIMENTAL PROGRAMME

This thesis is mainly concerned with the structure and regulation of the *A.nidulans* nitrate transporter, but includes some related aspects of nitrate assimilation.

1. 6. 1. Structure of the nitrate transporter

The *crnA* gene, thought to encode a nitrate transporter (1.4.2.) is located adjacent to the *niiA* gene (1.4.6.) but its exact position is uncertain. Complementation analysis (Greaves, 1990) suggest that the control region of the *crnA* gene lies between the *Sma*I and *Eco*RI sites (Fig 2.2). To investigate this and localise the precise position of the coding sequences it was thought necessary to sequence the whole region of DNA between the *Nru*I and *Eco*RI sites (Fig 2.2). Isolation from a library, and susequent sequencing, of the *crnA* cDNA would aid in the identification of any introns and the 5' and 3' ends of the gene. Confirmation of these sites

could be achieved by performing either S1 nuclease analysis or by primer extension.

The AA sequence encoded by the *crnA* gene can be deduced from the nucleotide sequence. A study should then reveal whether the protein contains any motifs common to other transporters (1.5.0.). Additionally, comparisons with available sequences of other transport proteins should highlight any areas of similarity.

1. 6. 2. Regulation of the nitrate transporter

Putative control regions identified from the sequence data (1.2.1) could be cloned into a plasmid harbouring a reporter gene, such as the *lacZ* system (Silhavy and Beckwith, 1985). Once transformed into various fungal strains mutant in any gene involved in the nitrate assimilation pathway, a simple enzyme extraction and spectrophotometric test would reveal which genes exert a regulatory effect on *crnA* gene expression. An alternative and proposed method for investigating the regulation of expression of the *crnA* gene is to examine *crnA* messenger RNA levels in the wild type strain under various growth conditions and nitrogen sources. Mutant strains of the various structural genes and regulatory genes involved with nitrate assimilation will be analysed in a similar way. Differential signals after S1 nuclease digestion of a *crnA* DNA-mRNA hybrid, using mRNA from the various mutant strains, would be a further approach to investigate the mode of regulation of *crnA* gene expression.

Upstream sequence comparisons of the *crnA* gene with the *niiA* and *niaD* genes will hopefully reveal the existence of putative *cis*-acting elements (1.2.1).

1. 6. 3. Regulation of *niiA* and *niaD* gene expression

Analysis of the regulation of *niiA* and *niaD* gene expression will be carried out in a similar way to that of the *crnA* gene. Aspects of the regulation of expression of these two genes are already known (Scazzocchio and Arst, 1989), but confirmation is required of the level of regulation (1.3.0.) as is the role played by the NR enzyme in autoregulation (1.4.7.).

It is presumed that nitrate assimilation in *A.nidulans* will be similar to that of its close relatives, *A.oryzae* and *A.niger*. Therefore it is proposed to examine the *niiA-niaD* intergenic regions of *A.oryzae* and *A.niger* to determine any areas of similarity with the corresponding regions of the *A.nidulans crnA, niiA* and *niaD* genes. This will hopefully identify genuine *cis*-acting regulatory sequences (1.2.1).

1. 6. 4. Heterologous expression in *A.nidulans*

The *nit*-2 gene of *N.crassa*, conferring nitrogen metabolite repression, has been shown to be expressed in the *A.nidulans* strain, *areA*19 (1.4.7.). The *nit*-4 gene of *N.crassa* is the equivalent of the *nirA* gene of *A.nidulans*, responsible for the induction of the nitrate assimilatory structural genes. It has been isolated on a plasmid, pNIT-4b (Fu *et al*, 1989) allowing the introduction, by a gene-mediated transformation system, into the *A. nidulans* strain mutant at the *nirA*1 locus. Growth tests on nitrate and NR assays of the transformants would indicate whether the *nit*-4 gene is able to substitute for the *nirA* gene of *A.nidulans*.

Chapter 2

MATERIALS AND METHODS

2. 1. 0. GENERAL

2. 1. 1. General Suppliers

All chemicals were of at least analytical grade and obtained from BDH or Sigma Chemical Co. unless otherwise stated. The radioisotope, α-^{32}P dCTP was supplied by ICN Biomedicals Inc.. Restriction enzymes and DNA modifying enzymes were purchased from Northumbria Biologicals Limited p.l.c. or Boehringer Corporation Limited. Novozym 234 was provided by NovoBiolabs, NovoIndustries.

2. 1. 2. Equipment

Ultracentrifugation was performed using a Beckman L2-65B. The Sorval RC-5C or RC-5B models (Du Pont) were used in conjunction with the Wifug 500E for medium and low speed centrifugation of large volumes. Micro-centrifugation was conducted in an Eppendorf centrifuge 5415 or a MSE microcentaur.

All absorbance readings were measured using a Pye Unicam SP6-550 UV/VIS Spectrophotometer. The pH meter was also supplied by Pye Unicam, model 292 Mk 2.

Lyophilisation of small volumes and removal of ethanol from nucleic acid samples was carried out using the Edwards Pirani 501 freeze drier.

Large and medium agarose gel electrophoresis was performed using the Bethesda Research Laboratories Co. (BRL) models H4 and H5 respectively, and the International Biotechnologies Inc. (IBI) model QSH was utilised for mini gel eletrophoresis. Power was supplied via the ATTA AE3121 or ST1082 models or the BRL 100 pack depending on the voltage required.

A UVP transilluminator TM40 allowed the photography of agarose gels with a Tominon polaroid MP-4 camera. X-rays were developed automatically with the Fuji RGII film processor.

A Bassaire A5HB laminar flow cabinet was available for sterile work. Batch cultures were grown in a New Brunswick model G25 shaking incubator.

A BBC masterseries computer was used to gain access to the software "analyseq" and "analysep" on the University of St.Andrews vax in addition to the UWGCG package on the Daresbury vax. Further analysis was undertaken with the DNA Strider 1.1 package on the Apple Macintosh Plus computer.

2. 1. 3. Media

2. 1. 3. a. Fungal (Modified from Cove, 1966).

Complete media (CM):	10 g	D-glucose
	1 g	yeast extract (lab m)
	1 g	casein hydrolysate (lab m)
	2 g	mycological peptone (lab m)
	25 ml	salt solution
	1 ml	vitamin solution
	1 ml	trace elements solution

Made up to 1 L and adjusted to pH 6.5 with 5 M NaOH. If required, agar was added to a final concentration of 1.2% (w/v).

Minimal media (MM):	10 g	D-glucose
	25 ml	salt solution
	1 ml	vitamin solution
	1 ml	trace element solution

Made up to 1 L and adjusted to pH 6.5 with 5 M NaOH. If required, agar was added to a final concentration of 1.2%.

Salt solution :

20.8 g KCl
20.8 g $MgSO_4 .7H_2O$
60.8 g KH_2PO_4 per L.

Vitamin solution :

1.5 g Aneurine
2.5 g Biotin
2.5 g Nicotinic acid
2.0 g Choline.HCl
0.8 g *p*-amino benzoic acid (PABA)
1.0 g Pyrodoxine.HCl
2.5 g Riboflavin
2.0 g Pantothenate (Ca salt) per L.

Trace element solution :

1.2 g $Na_2Mo_7O_{24}.2H_2O$
11.2 g H_3BO_4
1.6 g $CoCl.6H_2O$
1.6 g $CuSO_4.5H_2O$
50.0 g EDTA
5.0 g $MnCl_2.4H_2O$
22.0 g $ZnCl_2.7H_2O$

Made up to 1 L, heated to boiling, cooled to 60° C and adjusted to pH 6.5 with 1 M KOH (Hunter *et al*, 1950).

Fungal growth requirements were made to 100 x final concentration (in a 100 ml volume) and added to the media accordingly :

5.0 g L-methionine

7.5 g adenine

4.2 g arginine.HCl

0.1 g biotin

0.2 g putrescine

1.0 g pyridoxine.HCl

0.07 g PABA

0.1 g pantothenate (Ca salt)

The AAs; proline and urea and the nitrogen sources; glutamic acid (sodium salt), sodium nitrate, sodium nitrite and ammonium tartrate were added to a final concentration of 10 mM for the nitrogen ion.

2. 1. 3. b Bacterial (Sambrook *et al*, 1989)

Luria Broth (LB): 10 g NaCl

10 g Tryptone

5 g Yeast extract

Made up to 1 L and adjusted to pH 7.5 with 5 M NaOH. If required, agar (lab m) was added to a final concentration of 1.2%.

SOC : 2.0% Tryptone

0.5% Yeast extract

10.0 mM NaCl

2.5 mM KCl

10.0 mM $MgCl_2$

20.0 mM D-Glucose

10.0 mM $MgSO_4$

Made up to 1 L and adjusted to pH 7.5 with 5 M NaOH.

All bacteriophage work was carried out with NZYDT media (BRL). Liquid media consisted of 21 g of NZYDT in 1 L. If required, agar was added to a final concentration of 1.2%.

Media was sterilised by autoclaving for 15 minutes at 15 pounds per square inch (p.s.i.). Small volumes or heat labile solutions were passed through a 0.2 μm Dynaguard filter (Microgon).

2. 1. 4. Strains

2. 1. 4. a. Fungal

Gene symbols are as defined by Clutterbuck (1984). The *A.nidulans* strain *biA*1 was used as wild type. Strains supplied by Dr. J. Clutterbuck include G034; *biA*1 *argB*2, G0228; *biA*1 *nirA*^{c}1, G059; *biA*1 *cnxE*17, and G834; *yA*2 *pyroA*4 *nirA*1. The strains B5; *biA*1 Δ506, B125; *biA*1 *niaD*26, B468; *yA*1 *puA*2 Δ509, B344; *yA*1 *wA*3 *niaD*10, B466; *yA*1 *puA*2 Δ507, B351; *biA*1 *niaD*20, B397; *biA*1 *puA*2 *niaD*40, B352; *biA*1 *niaD*21, B348; *biA*1 *niaD*16, B556; *pantoB*100 *fwA*1 *niaD*118, B400; *biA*1 *puA*2 *niaD*42, B286; *yA*1 *biA*1 *niaD*5, B366; *biA*1 *niiA*17, B371; *biA*1 *niaD*25, B34; *pabaA*1 *fwA*1 *niaD*35, B564; *pantoB*100 *fwA*1 *niaD*124, B117; *pabaA*1 *fwA*1 *niaD*8, B361; *biA*1 *niaD*54, B552; *pantoB*100 *fwA*1 *niaD*115, B261; *yA*1 *pantoB*100 *fwA*1 *niaD*123, B269; *biA*1 *fwA*1 *niaD*28, B246; *yA*1 *puA*2 *niaD*127, B320; *biA*1 *niaD*31, B557; *pantoB*100 *fwA*1 *niaD*119, B558; *pantoB*100 *fwA*1 *niaD*120, B281; *pabaA*1 *niaD*2, B272; *pantoB*100 *niaD*29, B349; *biA*1 *niaD*17, B370; *biA*1 *niaD*24, B152; *biA*1 *niaD*34, B279; *yA*1 *biA*1 *niaD*1, B598; *biA*1 *niaD*171, B597; *biA*1 *niaD*170 and B9; *yA*1 *adE*20 *niaD*15 were supplied by Dr. B. Tomsett. Other strains used include N1458 *yA*2 *biA*1 *crnA*1 (provided by Professor H.N. Arst) and MH205; *biA*1 *niiA*4 *areA*19 and MH837; *biA*1 *xprD*1 (provided by Professor M.J. Hynes).

2. 1. 4. b Bacterial

Bacterial strains used for the maintenance and propagation of plasmids were:

E.coli JM101	*supE thi*Δ (*lac-proAB*) F' [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* Δ*M*15] (Sambrook *et al*, 1989)
E.coli DH5	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1 (Sambrook *et al*, 1989).

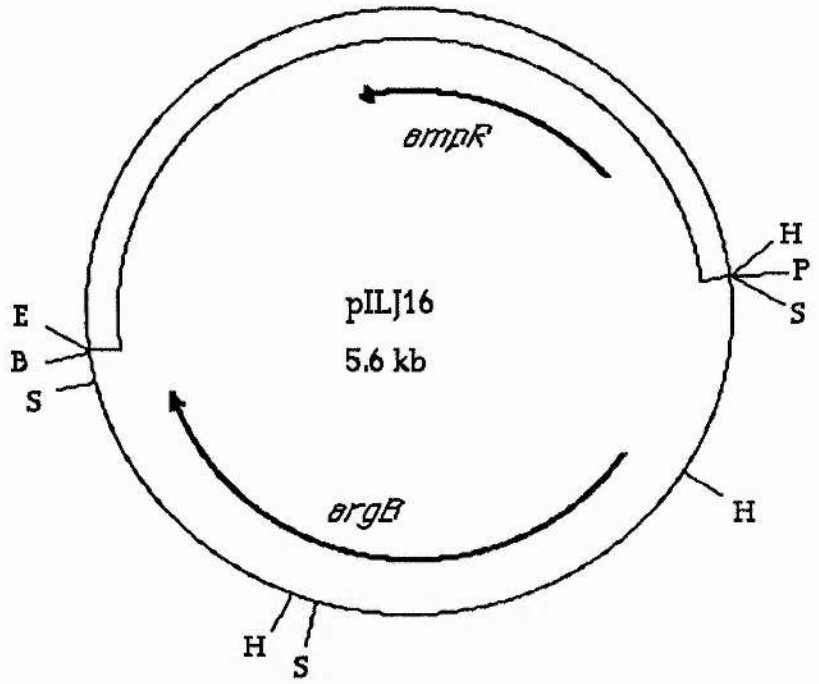


Figure 2.1 The vector pILJ16, used to subclone restriction endonuclease fragments of the nitrate gene cluster. It consists of the *A.nidulans argB* gene cloned into the plasmid pUC18, containing the *ampR* gene. Restriction enzyme symbols are E; *Eco*RI, H; *Hind*III, S; *Sal*I, P, *Pst*I and B; *BamH*I.

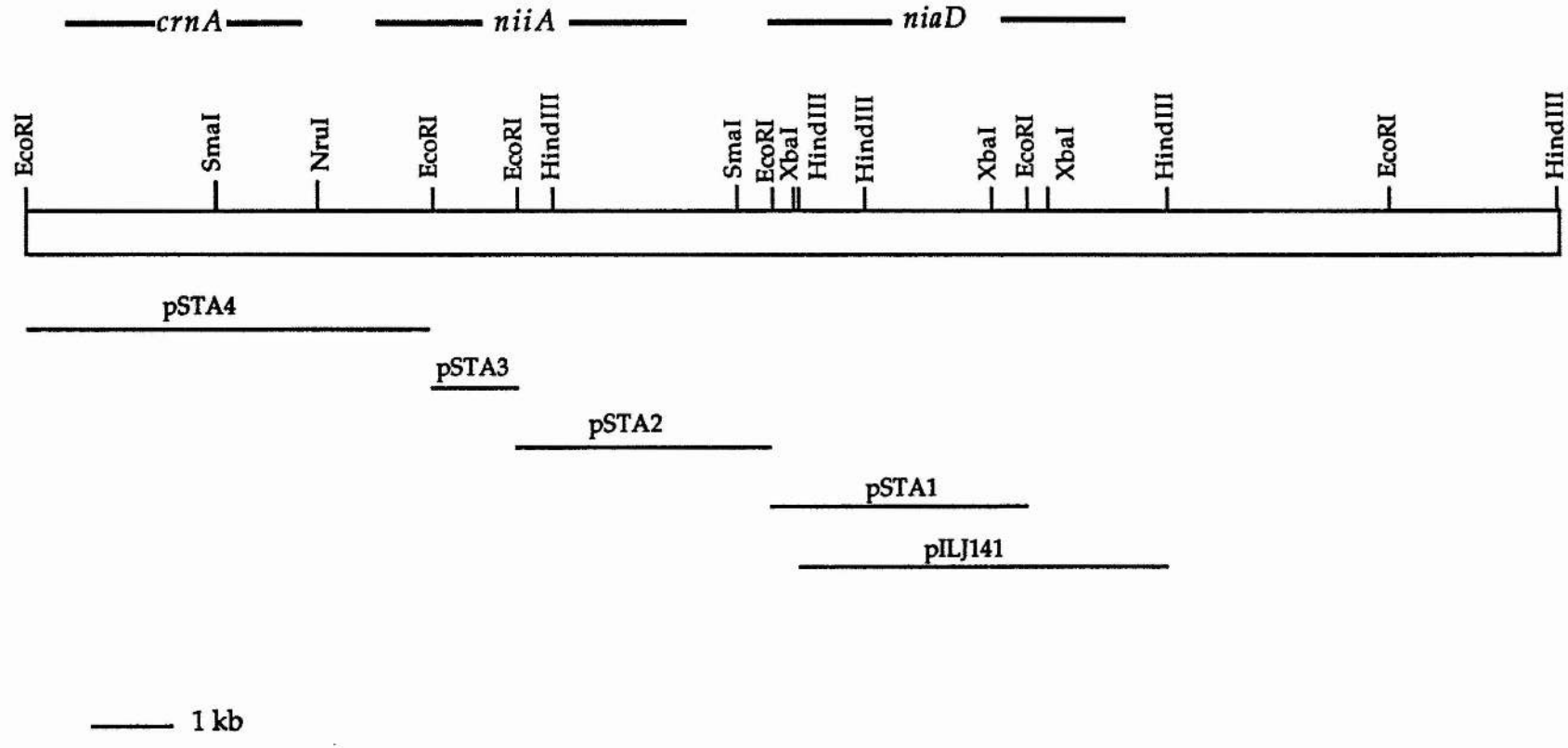


Figure 2.2 Restriction map of the *A.nidulans* nitrate gene cluster, indicating the extent of each of the subclones.

The host strain for bacteriophage λgt10 was:

E.coli NM538 *supF hsdR trpR lacY* (Sambrook *et al*, 1989)

These strains were stored in glycerol at -70° C.

2. 1. 4. c Vectors

The *A.nidulans* cDNA library was cloned in the bacteriophage λgt10 (Sambrook *et al*, 1989). The vector pILJ16 (Fig 2.1), a derivative of pUC18 (Sambrook *et al*, 1989) with the *A.nidulans argB* gene, for arginine utilisation, had been used to clone fragments of the nitrate cluster genes (Johnstone *et al*, 1990). These constructs were named pILJ141 and pSTA1, 2, 3 and 4 (Fig 2.2). The vectors, pNIT4 and pN4E2A, housing all or part of the *N.crassa nit-4* gene (1.6.4.) were obtained from Dr. G.A. Marzluf (Fig 2.3). The cloning vehicle pUC18 was also required for work in this thesis.

2. 1. 5. General Techniques

2. 1. 5. a Tris saturation of phenol

A quantity of 500 ml crystalline phenol was melted at 60° C and 0.1% hydroxyquinoline added in order to prevent oxidation and to colour the phenol lemon yellow. Extractions with 500 ml of 1 M Tris HCl pH 8.0 were conducted in a separating funnel until the aqueous phase became clear and the pH of the phenol was pH 8.0. Finally, two extractions with 0.1 M Tris HCl pH 8.0 were performed before the phenol was stored at 4° C, overlayed with 0.1 M Tris HCl pH 8.0.

2. 1. 5. b Preparation of phenol : chloroform

The chloroform contained 4% (v/v) iso amyl alcohol (IAA) to permit separation of the phases. The phenol and chloroform were mixed in the appropriate ratio well before use to allow clearing.

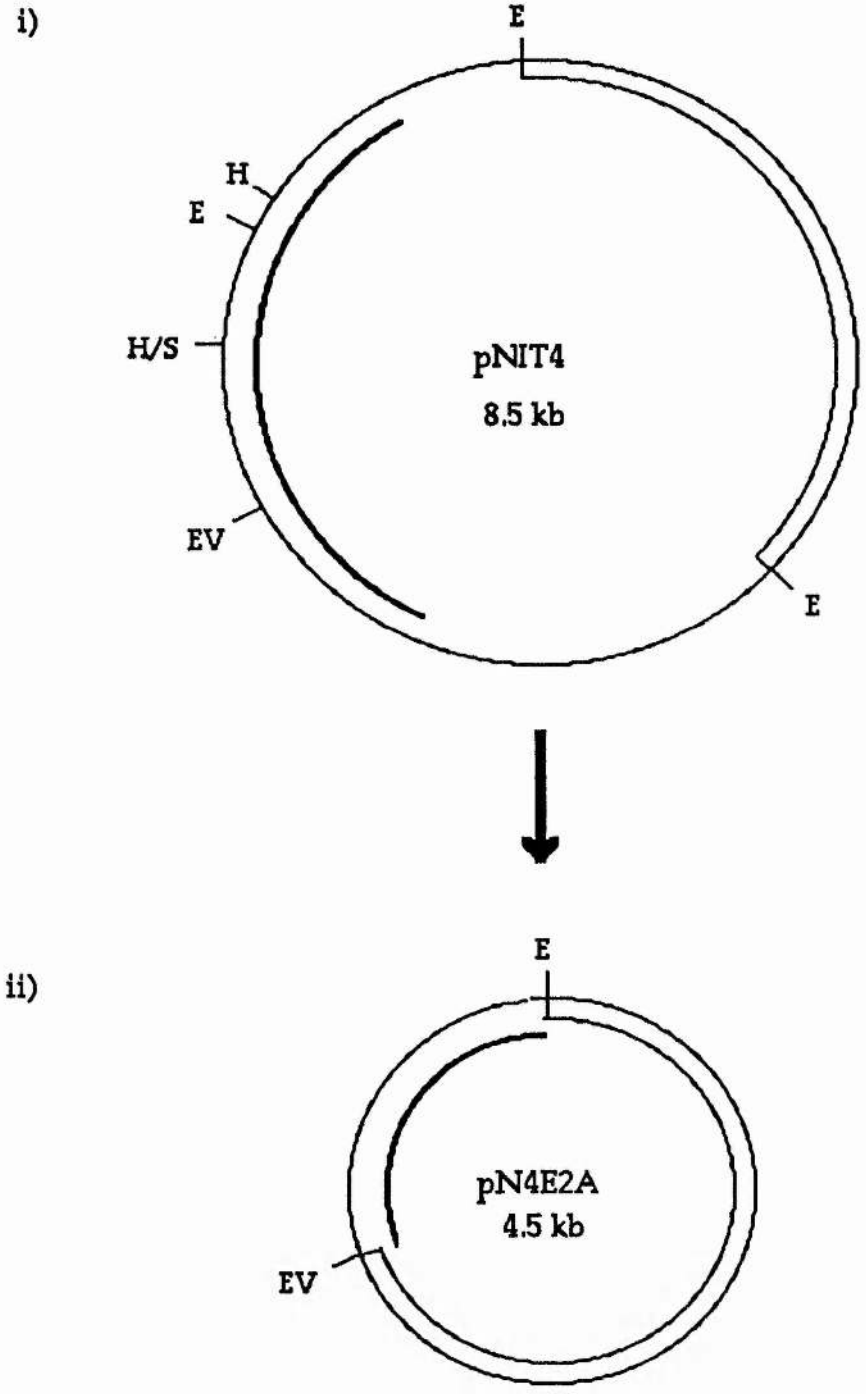


Figure 2.3 i) The vector pNIT-4, containing the *N.crassa nit-4* gene, cloned into a bluescript II KS plasmid. ii) The pN4E2A vector contains the 1.3 kb *EcoRI-EcoRV* fragment from the *nit-4* gene. The solid black line represents *N.crassa* sequences. Plasmid sequences are represented with a double line. Restriction enzyme sites are abbreviated as E; *EcoRI*, EV; *EcoRV*, H; *HindIII* and S; *SalI*. (Fu and Marzluf, 1989).

2. 1. 5. c. Deionisation of formamide

To every 100 ml of formamide to be deionised was added 25 g of amberlite monobed resin. The mixture was stirred in the dark for 30 min and filtered twice through Whatman No. 1. Formamide was stored in the dark at -20°C.

2. 2. 0. FUNGAL TECHNIQUES

2. 2. 1. Generation of double mutant strains (Clutterbuck, 1974)

An *A.nidulans* sexual cross between two strains each carrying a genetic marker of interest results in many progeny, some of which may carry both markers.

Strains were chosen with different coloured conidia, either green, yellow or white, to provide an initial, visual test for the success of the cross. A loop of conidia from each strain was mixed in the centre of an MM agar plate containing 5mM ammonium. A loop of CM and any supplements required by either strain was added to the conidia. The suspension was then spread in six directions over the surface of the agar. "Sellotape" was used to seal the plastic, 9 cm petri-dish plate (Sterilin) to prevent the entry of oxygen and to encourage the cross to occur. The plate was incubated in the inverted position at 37° C for at least five days.

After this period a binocular microscope was used to search for mature, black, spherical cleistothecia in the proximity of conidia from both strains. It is more likely that the cleistothecia in this region result from a sexual cross. If no cleistothecia were seen or if they appear red, incubation was continued to allow their maturation.

Cleistothecia were picked with a sterile needle and rolled on 2% agar to remove any conidia that may be present. The ascospores were released by squeezing the cleistothecia between the needle and the wall of a

universal containing 10 ml of saline tween (0.9% NaCl, 0.01% Tween 80). The ascospore-containing saline tween was vortexed and 200 µl spread over a CM plate (containing any of the growth requirements of either parent) before incubating at 37° C for a period of 48 h.

Progeny from a single cleistothecia of both parental colours is indicative of a successful cross. Phenotypic tests for the marker(s) of interest were conducted by replica plating a number of progeny onto MM with and without the metabolic supplement in question. Those progeny that only grow on the supplemented plate were presumed to be auxotrophic for that supplement and to have the genetic mutation of interest.

2. 2. 2. Fungal transformation (Johnstone *et al*, 1985 and Campbell *et al*,1989)

Fresh conidia (7 days old) from a 9 cm CM petri dish (10^8-10^9) were seeded into a 1 L unbaffled conical flask containing 200 ml of MM with 5 mM ammonium and any supplements required by that particular strain. The culture was grown for 12-15 h at 37° C, with 300 rpm orbital shaking to prevent clumping of the mycelia. Harvesting was by filtration through a sterile, double layer of muslin into a Buchner flask.

Mycelia was washed thoroughly with 500 ml of solution A (0.8 M $MgSO_4.7H_2O$, 10 mM $NaPO_4$ buffer pH 5.8) and placed into a universal to which 5 mg ml^{-1} of Novozym 234 in 20 ml of solution A was added. The suspension was incubated at 30° C with gentle shaking, for between 45 min and 2 h, or until protoplasting was complete. The viscous cellular suspension changes to a more liquid consistency.

Cellular debris is spun out at 3000 rpm for 10 min leaving the protoplasts suspended in the 0.8 M $MgSO_4$ above. This was divided into four plastic universals (Sterilin) and diluted out at least four times with

solution B (1.2 M sorbitol, 50 mM $CaCl_2$ and 10 mM Tris HCl pH 7.5). Protoplasts were spun out at 2000 rpm for 5 min. The pellets were resuspended in 5 ml of solution B, pooled and centrifuged again at 2000 rpm. Washes were repeated a further three times, to remove residual Novozym, before counting the protoplasts using a haemacytometer.

It is desirable for protoplasts to be at a concentration of 10^7 100 μl^{-1}. Up to 20 µg of the DNA to be transformed (this may be introduced via one or more plasmids) was added to the protoplasts in a volume not exceeding 20 µl. The mixture was allowed to stand for 20 min at room temperature. Subsequently the DNA was encouraged to cross the protoplast membrane by plunging the transformation mixture into 1 ml of solution C (50% polyethylene glycol (PEG) 4000, 50 mM $CaCl_2$, 10 mM Tris HCl pH 7.5) in a plastic universal. After 5-10 min and gentle shaking the PEG was diluted out by adding 5 ml of solution B. The contents of the universal were poured into 200 ml of cool MM agar with 1.2 M sorbitol to act as an osmotic stabiliser for the protoplasts. Any necessary metabolic requirements were added, omitting the supplement that only non-transformed protoplasts will require. Thin plates are poured to ensure distribution of the protoplasts.

A series of protoplast dilutions (generally 10^7, 10^5 and 10^3 ml^{-1}) in 100 µl of solution B were treated in the same way as the concentrated protoplasts undergoing transformation. However, no DNA was added to these and they were allowed to grow on media with all the supplements required by the parent strain. This allows an assessment of the ability of the protoplasts to regenerate. Novozym 234 and PEG can inhibit protoplast regeneration. Negative controls were constructed in which the same protoplast dilutions, treated in an identical fashion, are added to the same media as those protoplasts undergoing transformation. Any fungi

growing on these plates, in common with the transformant plates, were presumed to be contaminating strains.

Plates were incubated at 37° C in the inverted position for 5 days or until transformants regenerated.

2. 2. 3. Storage of fungal strains

All fungal strains were stored in silica, remaining viable for up to two years in a dessicated environment at 4° C.

Conidia, from the desired strain, were suspended in 1 ml of 5% (w/v) Marvel (Boots p.l.c.) in a sterile 1.5 ml eppendorf tube. Approximately 500 µl of this was transferred, dropwise, to a small vial containing silica (mesh 60-120) on ice. The silica was allowed to dry and any lumps disrupted before being stored as above.

2. 2. 4. Isolation of DNA (Kolar *et al*, 1988)

A conidial suspension of the appropriate fungal strain was used to inoculate a 1 L conical flask containing 200 ml MM with 10 mM ammonium supplemented with 0.5% yeast extract. Incubation proceeded at 37° C, with 250 rpm orbital shaking for 20 h after which harvesting of mycelia took place by filtration through a sterile, double layer of muslin.

The mycelia was squeezed and pressed between paper towels to remove excess liquid before freezing in liquid nitrogen. Frozen mycelia was ground to fine powder under liquid nitrogen, in a mortar and pestle, and suspended in 10 ml of TEN buffer (0.2 M Tris HCL pH 8.5, 0.05 M EDTA, 0.25 M NaCl, 48 mg ml^{-1} *p*-amino salicylic acid (PAS) and 8 mg ml^{-1} triisopropyl napthalene sulphonic acid (TNS) (Kodak)) in a sterile 30 ml polyallomer centrifuge tube (Du Pont). This buffer was freshly prepared and left to stand on ice to allow the precipitate to settle.

The reaction mixture was extracted with an equal volume of phenol pH 7.9 : chloroform (1 : 1), agitated gently and centrifuged for 5 min at 4000 rpm to separate the phases. For higher yields the organic phase was back extracted. A further two phenol : chloroform extractions were performed with the supernatants.

DNA was precipitated by adding sodium acetate (NaOAc) to 0.3 M followed by 2-3 volumes of cold (-20° C) 96% ethanol. Gentle agitation promoted agglutination of the DNA which was then collected with a sterile, hooked pasteur pipette. The DNA was washed with 70% ethanol, vacuum dried and resuspended in a suitable volume of TE (10 mM Tris HCl pH8.0, 1 mM EDTA). The quantity of DNA recovered, generally 500 µg, was measured spectrophotometrically, given an OD of 1.0 at 260 nm is equivalent to 50 µg ml^{-1} of DNA. The A_{260}/A_{280} ratio provided a measure of the purity of the DNA, 1.8 being indicative of a pure preparation.

Smaller quantities of DNA could be prepared using a scaled down version of this method. Mycelia were grown in 10 ml of media in a universal for a period of 48 h.

2. 2. 5. Isolation of RNA (Cathala *et al*, 1983)

A.nidulans strains for RNA preparations were grown at 30° C, with 250 rpm orbital shaking for 12 h. If a switch to a second nitrogen source was required, incubation proceded for a further 4 h unless otherwise stated. Cultures to be used for the temporal investigation of *crnA* gene expression were grown at 37° C for an initial period of either 10, 15, 20 or 30 h. Nitrate was subsequently added to 10 mM and cultures incubated as above for a further 100 min. Harvesting of mycelia was as described previously (2.1.4).

Approximately 1 g of mycelia, ground to a fine powder under liquid nitrogen in a mortar and pestle, was poured into 3 ml of stirring lysis

buffer (5 M GuSCN (Fluka), 50 mM Tris HCl pH7.5, 10 mM EDTA and 10% (v/v) of ß-mercaptoethanol is added immediately before use). The mycelial suspension was homogenised by passing four times through a 0.8 mm needle and once through a 0.6 mm needle. To the continuously stirring suspension was added 15 ml of 4 M LiCl and the mixture poured into a 30 ml Corex tube (Du Pont). The reaction was stood for 20 min at room temperature followed by over night incubation at 4° C.

Cellular debris was pelleted at 2500 rpm for 5 min in a swing-out HB4 rotor (Du Pont) at 4° C. Centrifugation of the supernatant proceeded at 8000 rpm for 90 min and the resulting pellet homogenised in 10 ml of 3 M LiCl through a 0.8 mm needle three times and a 0.6 mm needle twice. This suspension was spun out for 60 min at 8000 rpm and the pellet resolubilised in 1 ml of TESDS (10 mM Tris HCl pH 7.5, 1 mM EDTA and 0.1% sodium dodecyl sulphate (SDS)) by passing through a 0.8 mm needle until homogeneous.

The RNA was purified by extracting with phenol : chloroform (1 : 1) until no interface remained. To increase yields, the organic phase from the first extraction was re-extracted with 0.5 ml TESDS. Finally, RNA was ethanol precipitated as described previously (2.1.4).

Vacuum dried pellets were resuspended in DEP water to allow the A_{260} and A_{280} OD measurements to be recorded. Given that 40 µg ml^{-1} of RNA is equivalent to an OD A_{260} of 1.0 the prepared RNA was aliquoted into 100 µg samples, lyophilised and stored at -70° C. The average yield was approximately 600 µg. The $^{A260}/_{A280}$ ratio provided a measurement of the RNA purity, 1.8 being ideal.

All glassware was baked at 200° C for 8 h. Lysis buffer, prepared by dissolving the GuSCN at 50° C and adding the Tris and EDTA was filtered through a 0.4 µm millipore filter. All other solutions were autoclaved with 0.1% DEP and all equipment was autoclaved thoroughly.

2. 3. 0. ENZYME ASSAYS

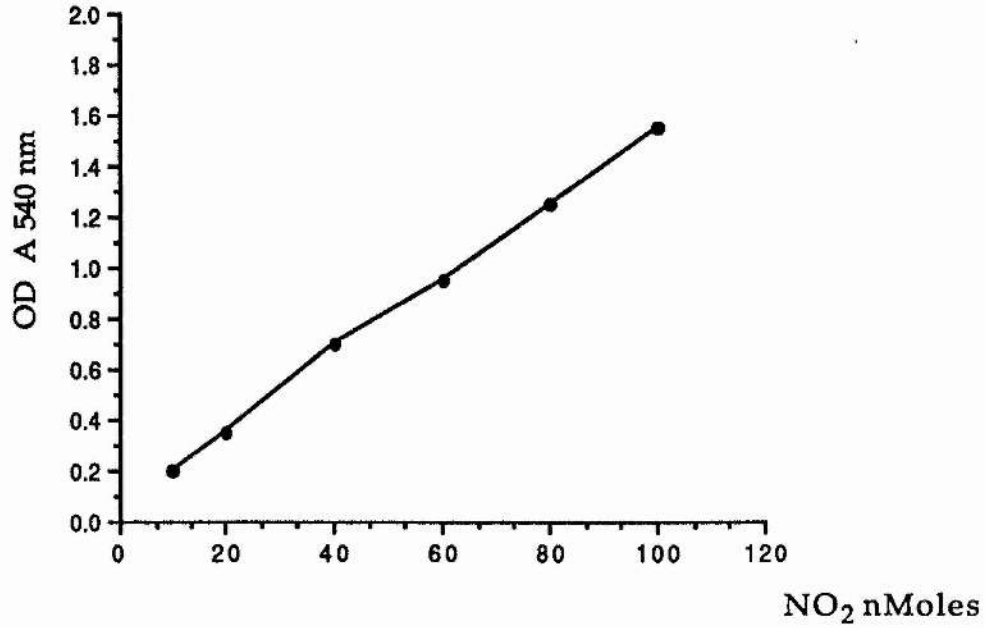
2. 3. 1. Nitrate reductase (Cove, 1966)

Mycelia were grown at 26° C for 18 h at 300 rpm orbital shaking in MM with the appropriate nitrogen source and supplements. A non-nitrate utilising strain may be initially grown in ammonium and transferred to MM with 10 mM nitrate for 4 h. Mycelia was harvested as described previously (2.1.4), ensuring the removal of all excess liquid by squeezing the mycelia between paper towels.

A 0.5 g aliquot of mycelia was ground to a fine powder under liquid nitrogen with a mortar and pestle. The mycelia were suspended in 5 ml of 100 mM sodium phosphate buffer pH 7.2, in a 30 ml centrifuge tube, and gently shaken before centrifuging at 15,000 rpm for 15 min at 4° C in a SS34, fixed angle rotor (Du Pont). Approximately 10 µl of the supernatant was added, in triplicate, to test tubes and the volume adjusted to 600 µl with the sodium phosphate buffer. The following solutions were then added to the extract, mixed and incubated at 25° C for 20 min: 100 µl of 50 mM Na_2SO_3, 100 µl of 100 µM FAD, 100 µl of 2 mM NADPH and 100 µl of $NaNO_3$.

To complete the reaction 1 ml of sulphonilamide solution (1% sulphonilamide in 3 M HCl) and 1 ml of 0.2% N-1-Napthylethylenediamine dihydrochloride (NED) were added to each test tube, vortexed and incubated in the dark for 10 min. The absorbance was read at 540 nm and converted to nanomoles of nitrite produced in twenty minutes from a standard curve (Fig 2.4). The specific activity of the enzyme was expressed as nanomoles of nitrite produced hr^{-1} mg $protein^{-1}$.

i)



ii)

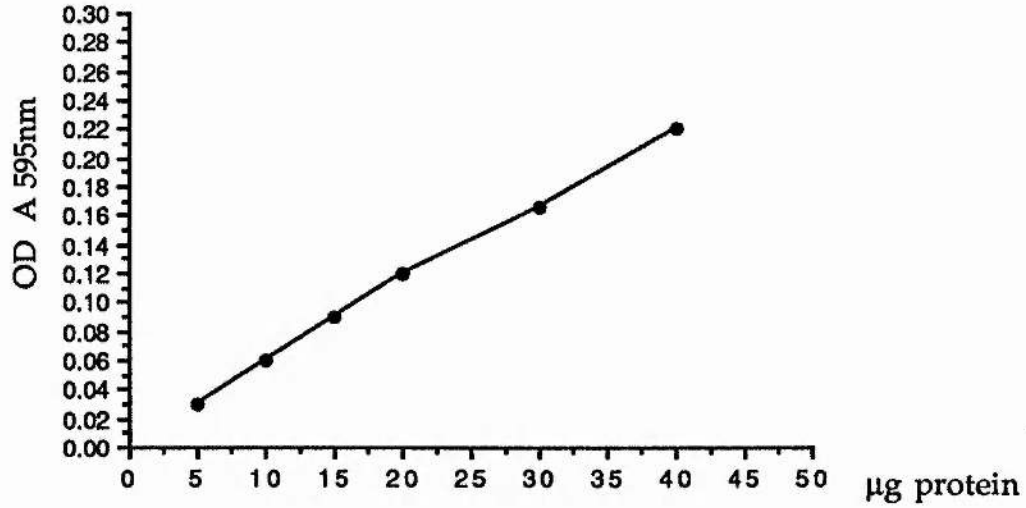


Figure 2.4 Standard curves of absorbance against; i) nitrate reductase activity expressed as nmoles of NO_2 synthesized and ii) micrograms of protein, as determined by the method of Bradford (1976).

2. 3. 2. Protein determination (Bradford, 1976)

Similar quantities of mycelial extracts as those used in the nitrate reductase determinations (2.3.1) were aliquoted in triplicate and vortexed in a test tube with 5 ml of Bradford's reagent (see below). The absorbance was measured immediately at 595 nm and converted to micrograms of protein using a standard curve (Fig 2.4).

Bradford's reagent was prepared by dissolving 0.1 g of Coomassie brilliant blue G in 50 ml of 96% ethanol. To this was added 96 ml of orthophosphoric acid and the solution made up to 1 L with distilled water. The reagent was filtered twice through Whatman no.1 filter paper and stored in the dark at 4° C for up to 60 days.

2. 4. 0. BACTERIAL TECHNIQUES

2. 4. 1. Preparation of competent cells (Hanahan, 1983)

A single bacterial colony was used to seed 10 ml of LB and the culture incubated at 37° C over night. The concentrated bacterial culture was subsequently used as an inoculum for 200 ml of LB in a baffled 1 L conical flask at a 100 x dilution.

When the OD_{600} of the bacterial culture had attained 0.2-0.3 U the cells were transferred to a sterile 500 ml polyallomer centrifuge tube and pelleted by centrifuging at 5000 rpm, 4° C in the GS3 rotor (Du Pont). The supernatant was discarded and the pellet resuspended gently in 50 ml of 100 mM $MgCl_2$. This protects transforming DNA from bacterial DNases (2.4.2). The cells were spun down as before and gently resuspended in 50 ml of 100 mM $CaCl_2$. The positively charged calcium ions act as an attractant for the negatively charged DNA during transformation (2.4.2). Incubation on ice for a period of 20 min to 1 h proceeded before

Incubation on ice for a period of 20 min to 1 h proceeded before centrifuging again at 5000 rpm for 5 min. Finally, the cells were resuspended in 10 ml of 100 mM $CaCl_2$ and 15% glycerol, aliquoted in 1 ml samples and stored at -70° C.

The competency of the cells, normally 10^6-10^7, was examined by performing transformations (2.4.2) with varying quantities of a plasmid housing a selectable marker (usually ampicillin).

2. 4. 2. Bacterial transformation

Competent cells (2.4.1) of the desired bacterial strain were thawed on wet ice. A quantity of DNA (enough to produce single bacterial colonies when transformed into the competent cells in question) was placed into a sterile 1.5 ml eppendorf tube with 100 µl of competent cells and the volume made up to 200 µl with 100 mM $CaCl_2$ solution. The mixed contents of the tube were incubated on ice for 20-30 min and heat shocked at 42° C for 2 min to encourage the DNA to enter the bacterial cell. SOC was added to 1.2 ml before incubating at 37° C for 60 min with gentle shaking to allow the expression of the *ampR* gene. A volume of 200 µl was spread, in duplicate, over the surface of an LB plate containing 100 µg ml^{-1} of ampicillin (Stock ampicillin was 20 mg ml^{-1} stored at -20° C for up to 3 months). Plates were incubated at 37° C over night in the inverted position.

Several plasmid vectors contain, in addition to the ampicillin resistance gene, the *E.coli lacZ* gene encoding ß-galactosidase. The transformation of such a vector into a *lac*$^-$ bacterial strain, e.g. JM101 (2.1.4.b), will yield blue colonies on LB plates containing 100 µg ml^{-1} ampicillin, 250 µg ml^{-1} isopropyl ß-D-thiogalactosidase (IPTG) and 250 µg ml^{-1} 5-Bromo-4-chloro-3-indolyl-ß-D-galactopyranoside (X-gal). The ampicillin selects for the transformed bacteria, IPTG induces the

the presence of ß-galactosidase (a blue colouration). This system can be used to investigate the success of a ligation (2.5.1.c). The multiple cloning site of the *lacZ* plasmids occurs within the *lacZ* gene. Therefore, any DNA fragment ligated into these vectors disrupts the *lacZ* gene and prevents ß-galactosidase activity. This results in white bacterial colonies, easily differentiated from the blue colonies housing vectors without inserts.

2. 4. 3. Alkaline lysis method for the preparation of plasmid DNA (Sambrook *et al*, 1989)

This method can be used to obtain pure plasmid DNA from a concentrated bacterial culture. The bacterial chromosomal DNA is selectively denatured with alkali under conditions where the plasmid DNA remains native. On neutralisation, chromosomal DNA becomes insoluble. Most protein and rRNA is removed by precipitation with SDS and high salt. The plasmid DNA can then be precipitated from the supernatant with ethanol.

The bacterial colonies of interest were grown over night in 2 ml of LB with 100 µg ml^{-1} of ampicillin at 37° C. Half of the culture was transferred to a sterile eppendorf tube and centrifuged for 1 min, the remainder was stored at 4° C. The supernatant was aspirated and the pellet resuspended by vortexing in 100 µl of ice cold lysis solution (50 mM glucose, 10 mM EDTA, 25 mM Tris HCl pH 8.0 and 4 mg ml^{-1} lysozyme). Incubation proceeded at room temperature for 5 min before 200 µl of freshly prepared ice cold solution B (0.2 M NaOH, 1% SDS) was added. The contents of the tube were mixed by inverting and left on ice for a further 5 min before a volume of 150 µl of ice cold solution C (Potassium acetate pH 4.8 made by adding 11.5 ml of glacial acetic acid to 60 ml of 5 M KOAc, and 28.5 ml of distilled water. The resulting solution is 3M with

respect to potassium and 5 M with respect to acetate.) was added. The mixture was vortexed gently for 10 s and again incubated on ice for 5 min.

The precipitate was pelleted by centrifugation at 4° C for 5 min and the supernatant transferred to a fresh, sterile eppendorf tube in which it was extracted with an equal volume of phenol : chloroform (1:1) (2.1.5.b). Two volumes of 96% ethanol were vortexed with the aqueous phase and the tube allowed to stand for 2 min at room temperature. The plasmid DNA was ethanol precipitated as described previously (2.2.4).

2. 4. 4. Large scale plasmid preparation (Sambrook *et al*, 1989)

This method follows the same principle as that of section 2.4.3. and is basically a scaled up version. A single bacterial colony containing the plasmid of interest was seeded into 10 ml of LB with 100 μg ml^{-1} ampicillin and grown over night at 37° C. A 2 ml volume of this culture was used to inoculate 200 ml of LB with 100 μg ml^{-1} ampicillin in a baffled 1 L conical flask. The culture was again grown over night at 37° C.

The bacterial culture was transferred to a sterile 500 ml centrifuge tube and the cells pelleted in the GS3 rotor at 8000 rpm for 10 min at 4° C. The supernatant was decanted and excess liquid removed before disrupting the pellet in 12.5 ml of solution A (50 mM glucose, 10 mM EDTA, 25 mM Tris HCl pH8.0). After 30 min incubation on ice, 20 ml of solution B (2.4.2) was added, the tube inverted to mix and left on ice for a further 5 min. Finally, 10 ml of solution C (2.4.2) was added, the mixture gently agitated and incubated on ice for 15 min.

The precipitate was pelleted by centrifugation in the GS3 rotor at 8000 rpm for 5 min at 4° C. The supernatant was passed through a double layer of sterile muslin into a second sterile 500 ml centrifuge tube and mixed with 0.6 x volume of cold (-20° C) propan-2-ol. Centrifugation at 8000 rpm for 5 min pelleted the plasmid DNA which was vacuum dried

and taken up in 4 ml of TE. To remove unwanted RNA the plasmid DNA was treated with 20 μg ml^{-1} of RNase A at 37° C for 1-2 h in a sterile 30 ml centrifuge tube.

To purify the DNA the 4 ml sample was applied to a "pack 500" Qiagen column (Qiagen Inc.). This consisted of anion-exchange resin and allowed the separation of DNA from proteins, polysaccharides and other contaminants by chromatography. The salt concentration of the DNA sample was adjusted to 850 mM by adding 880 μl of 5 M NaCl and the pH altered to 7.0 by adding 400 μl of 1 M 3-[N-morpholino]propanesulphonic acid (MOPS) buffer pH 7.0. These conditions allowed maximum adsorbtion of the DNA onto the column. Before applying the DNA sample, the column was equilibriated with 5 ml of buffer A (400 mM NaCl, 15% ethanol, 50 mM MOPS pH 7.0) forced through with a syringe at 2 ml min^{-1}. The plasmid solution was applied at a rate of 1 ml min^{-1} to the "pack 500" to ensure binding to the resin. The column was washed with 20 ml of solution C (1 M NaCl, 15% ethanol, 50 mM MOPS pH 7.0) at 2.5 ml min^{-1}. The higher salt concentration and faster flow rate removed all proteins and RNA that may have become adsorbed onto the column. Finally, the DNA was eluted with 5 ml of solution F (1.5 M NaCl, 15% ethanol, 50 mM MOPS pH 7.5), maximum recovery being achieved with a flow rate of 1 ml min^{-1}. The change of pH and increased salt concentration ensured that double stranded DNA was recovered. The salt concentration and pH may be altered in order to recover RNA or single stranded DNA, for example.

DNA was precipitated by the addition of 4 ml of cold (-20° C) propan-2-ol, followed by a 30 min incubation on ice. Centrifugation for 15 min in the SS34 rotor at 15,000 rpm, 4° C pelleted the DNA which was then washed in 70% ethanol and repelleted. After vacuum drying, the DNA

was taken up in an appropriate volume of TE, the quantity recovered being determined spectrophotometrically (2.2.4).

2. 4. 5. Large scale bacteriophage preparation (Davis *et al*, 1986 and Zabarovsky and Turina, 1988)

The method of bacteriophage DNA preparation exploits the biology of the virus. The bacteriophage adsorbs to receptors on the outer membrane of *E.coli*, encoded by the *lamB* gene, used to transport maltose. For this reason maltose is included in media used to grow bacteriophage. In addition, the inclusion of magnesium ions aids adsorbtion. After entering a host the bacteriophage replicates either via the lytic cycle or the lysogenic cycle. During lytic growth, important for DNA preparation, the viral DNA replicates many fold. Bacteriophage progeny are then assembled and the bacterial cell lyses, releasing the viral particles. An infected bacterial culture grown to lysis ensures a maximum concentration of bacteriophage DNA within the media. This removes the bacterial lysis step from the DNA preparation method, allowing direct precipitation of viral DNA by the addition of PEG and salt.

A small scale lysate from a single plaque was first generated. 100 µl of host cells, *E.coli* NM538 (2.1.4.b), grown to an OD_{600} of 0.1-0.2 are incubated at 37° C with the desired plaque (removed from a plate in a plug of agar using a pasteur pipette), for 15 min with gentle shaking. The cells and phage were transferred to 4 ml of NZYDT media (or LB supplemented with 0.2% maltose) in a sterile 30 ml centrifuge tube and incubated at 37° C with 300 rpm orbital shaking until lysis occurs, usually after 4-6 h. A lysed bacterial culture loses its "silky" sheen and the bacterial debris forms a precipitate within the culture. A few drops of chloroform can be added to lyse any remaining bacteria. Cellular debris

was removed by spinning at 4000 rpm for 5 min and the lysate stored at 4° C with a few drops of chloroform to prevent further bacterial growth.

It was necessary to titre the lysate (2.5.9.) to ensure that the correct number of plaque forming units (pfu) are added to the cells for a large scale DNA preparation.

To prepare λgt10 DNA a second, larger liquid lysate with a titre of at least 5×10^8 pfu ml^{-1} was generated. A 100 ml volume of NZYDT media in a 500 ml conical flask was inoculated with a single colony of NM538 and incubated overnight at 37° C, with 300 rpm orbital shaking. The OD_{600} of the culture was measured and an aliquot containing 10^6 cells removed (given that an OD_{600} of 1.0 is equivalent to 8×10^8 cells ml^{-1}). Centrifugation at 4000 rpm for 10 min at room temperature pelleted the cells which were then resuspended in 3 ml of TM buffer (50 mM Tris HCl pH 7.4, 10 mM $MgSO_4$).

Bacteriophage from the small scale lysate were added to the cells. The number required to obtained a large scale lysate with a good titre was critical, for λgt10, generally 5×10^7 pfu were required. The cell and bacteriophage were mixed in a 75 x 12 mm sterile test tube (Sterilin) and incubated for 20 min at 37° C, 100 rpm to allow adsorbtion of the bacteriophage.

A 1 L conical flask containing 200 ml of pre-warmed NZYDT media was inoculated with the contents of the test tube and incubated at 37° C, 300 rpm. Lysis should occur 9-12 h after inoculation but may be encouraged by the addition of a few drops of chloroform. If lysis occurs too early it is likely that the titre, measured as described (2.5.9), will be too low for efficient DNA yields. In this case lysis should be repeated and the cell : bacteriophage ratio increased. Conversely, if lysis fails, the cell : bacteriophage ratio should be decreased.

The lysate was decanted into a sterile 500 ml centrifuge tube and spun at 4000 rpm for 10 min at 4° C in the GS3 rotor to remove bacterial debris. The supernatant was transferred to a fresh tube, the salt concentration adjusted to 1 M, and 10% (w/v) PEG 6000 added. The contents were mixed and incubated on ice for 30 min. The phage DNA was pelleted using the GS3 rotor at 4000 rpm, 4° C for 10 min and resuspended in 5 ml of TE containing 10 mM $MgCl_2$. Extraction with an equal volume of choroform (the phases are separated at 4000 rpm for 5 min) was followed by ultracentrifugation in the SW64 rotor (Beckman) at 35,000 rpm for 20 min. The pellet was resuspended in 600 µl of TE with 10 mM $MgCl_2$ and any insoluble particles removed by a 2 min spin in a microcentrifuge. RNase A was added to the supernatant at a final concentration of 0.5 µg ml^{-1}. In addition, SDS to 0.1% and EDTA to 7.25 mM were added before incubating at 70° C for 5 min.

The DNA was purified by extracting once with an equal volume of phenol and twice with chloroform (the phases were separated as before) before ethanol precipitation (2.2.4). The yield, generally 500 µg was determined spectrophotometrically (2.2.4).

2. 5. 0. MOLECULAR BIOLOGY TECHNIQUES

2. 5. 1. DNA Restriction and modifying enzymes

2. 5. 1. a Restriction endonucleases

Restriction endonucleases are bacterial enzymes that cleave double stranded DNA. Their purpose is to destroy invading DNA (e.g. bacteriophage). The host DNA is protected from cleavage by modification of the recognition sites, usually by methylation. The enzymes are purified by the manufacturers (2.1.1) and supplied with a 10 x concentration assay buffer.

The temperature at which many endonucleases operate is 37° C although restriction with *Taq*I proceeds at 67° C. The addition of 100 μg ml^{-1} of nuclease free bovine serum albumin (BSA) is generally recommended, to provide optimal enzyme activity, but in most cases it was found to be unnecessary.

Reactions to be immediately analysed by gel electrophoresis were often carried out in a volume similar to the well size of the gel concerned. However, many digests involving larger quantities of DNA (20-100 μg) were conducted in larger volumes, 50-400 μl to aid digestion, and therefore ethanol precipitation was required before gel analysis (2.2.4).

In theory, 1 U of enzyme will digest 1 μg of λ DNA in 1 h under the correct assay conditions (Sambrook *et al*, 1989). However, in practice this was generally not the case due to insufficient purity of DNA and/or the presence of several restriction sites, therefore excess conditions (a combination of enzyme quantity and time) were employed. To ensure complete digestion, aliquots of the reaction were electrophoresed on a minigel at intervals and observed on the UV transilluminator. DNA refractory to digestion was purified by extraction with phenol : chloroform and ethanol precipitation (2.2.4).

When a double digest was necessary and the reaction conditions of the two enzymes were incompatible in terms of salt concentration, the enzyme with the lowest salt requirement was added to the reaction initially. Before the second enzyme was used either the salt concentration was adjusted or the reaction ethanol precipitated and the DNA resuspended in the second buffer.

The volume of enzyme included in a reaction should never be more than 10% of the final volume. Larger concentrations may result in inhibition of the reaction by the glycerol in which the enzyme is stored.

2. 5. 1. b. Calf intestinal phosphatase (CIP)

To ensure that vectors did not recircularise during a ligation (2.5.1.c) it was necessary to remove the 5' phosphate groups from the end of both DNA strands. Generally, 0.01 U of CIP was required to phosphatase 1 μg of DNA cut with an enzyme generating sticky ends. Incubation was sequential for a total of 1 h at 37° C in CIP buffer made at 10 x assay concentration (10 mM $ZnCl_2$, 10 mM $MgCl_2$, 100 mM Tris HCl pH 8.3).

The presence of CIP in a ligation will dephosphorylate insert DNA thereby preventing ligation (2.3.1.c). To ensure complete removal, the enzyme was denatured by heating to 75° C for 10 min. Protein was then removed by extraction with phenol : chloroform and the DNA was ethanol precipitated (2.2.4). In general, the CIP efficiency was at least 95%.

2. 3. 1. c. T4 DNA Ligase

A theoretical analysis of the parameters governing ligation was published by Dugaiczyk *et al*, 1975. The most significant factor in ligation is the concentration of DNA ends in the reaction mix, both the total concentration, denoted i, and j, the concentration of one end in the immediate proximity of the other end of the same DNA molecule. From this a mathematical equation was assembled:

$$MW = \left(\frac{51.1}{{}^{j}/_{i} \quad [DNA]} \right)^{2}$$

where MW is the molecular weight of the vector and [DNA] is the concentration of the vector (g L^{-1}).

It was found that a $j/_i$ ratio of less than 1 led to the formation of linear oligomers, in contrast a $j/_i$ ratio of greater than 2 promoted recircularisation. Therefore the appropriate $j/_i$ ratio is used for the type of ligation required.

The *crnA* cDNA *EcoR*I fragment isolated from a λgt10 library (2.5.9) was ligated into pUC18 linearised with *EcoR*I (2.3.1.a) and treated with CIP (2.3.1.b). The CIP treatment prevents recircularisation of the plasmid. Therefore the conditions immediately promote either concatenisation of the *crnA* cDNA fragment or its ligation into the vector.

The concentration of the vector in the ligation mix was 5 μg ml^{-1} giving a $j/_i$ ratio of 7, greatly favouring recircularisation. However, 95% of the vector was phosphatased, preventing recircularisation without the insertion of the cDNA fragment. The phosphate groups on the 5' ends of each strand of the cDNA allow the ligation reaction to occur. This results in two single stranded nicks, in the hybrid molecule, that are repaired after introduction into a competant bacterial cell (2.4.2). The vector : fragment molecular weight ratios were over a range of 1:1 to 1:10 to encourage fragment ligation into the vector.

Ligase buffer was made to 10 x assay concentration (500 mM Tris HCl pH 7.5, 100 mM $MgCl_2$, 100 mM DTT, 6.6 mM ATP) and stored at -20° C. The ligase was used at 0.1 U per reaction, typically in a volume of 20 μl and incubated at 22° C for 1 h for *EcoR*I ends. The ligation was arrested by heating at 68° C for 10 min before analysis by transformation into competent *E.coli* JM101 cells (2.4.2). X-gal and IPTG were included in the LB plates to select for bacteria carrying a plasmid with an insert (2.4.1).

Ligation controls included uncut plasmid, to test the transformation efficiency, cut plasmid with 0.1 U of ligase, to test the ligation efficiency and cut, CIP treated plasmid with 0.1 U of ligase to test the CIP efficiency.

2. 5. 2. Electrophoresis of nucleic acids (Sambrook *et al*, 1989)

2. 5. 2. a. Large DNA fragments (600 bp-25 kb)

DNA was electrophoresed through horizontal agarose gels. The concentration of the agarose was altered depending on the molecular weight of the DNA. Fragments of ~9 kb, and larger, were separated on 0.8% gels whereas 1.2% to 2% gels were used for fragments of between 3 kb and 600 bp. The size of the gel prepared, either BRL H4 or H5 models, was dependent on the number of samples and the duration of electrophoresis required. Low melting point gels for the isolation and recovery of DNA fragments were made with a comb having a single, wide well.

The gel was made with TBE buffer stored as a 10 x stock at room temperature (1 M Tris base, 1 M boric acid, 20 mM EDTA pH 8.0) and stained with 0.25 μg ml^{-1} EtBr, from a stock of 10 mg ml^{-1}. EtBr is a UV fluorescent dye that intercalates between stacked bp of DNA giving an immediate indication of the quantity of DNA in a gel when viewed on a transilluminator.

The agarose was dissolved by heating over a Bunsen burner and the gel poured when the temperature had cooled to 50° C. Low melting point gels, buffered in TAE (The 10 x stock solution was 0.4 M Tris base, 0.4 M sodium acetate, 10 mM EDTA), were allowed to solidify at 4° C.

Loading buffer (0.25 % Bromophenol blue, 0.25% Xylene cyanol FF (Kodak), 100 mM EDTA, and 30% (w/v) sucrose) was added to $^1/_{10}$ th volume of the DNA sample. The sucrose in the loading buffer has a high density, allowing the DNA sample to sink to the bottom of the well. EDTA was added to inhibit nuclease activity, and the dyes: bromophenol blue and xylene cyanol were included as a reference for the migration of the DNA. Their rate of migration alters depending on the concentration of the gel. Samples were introduced to individual wells once the running

buffer (1 x TBE) was in place and *Hind*III cut λDNA was used as molecular weight markers (Fig 2.5).

The running conditions employed were determined by convenience. Generally 15 V was required to allow the gel to run over night although minigels were electrophoresed at 100 V if a quick result was desired.

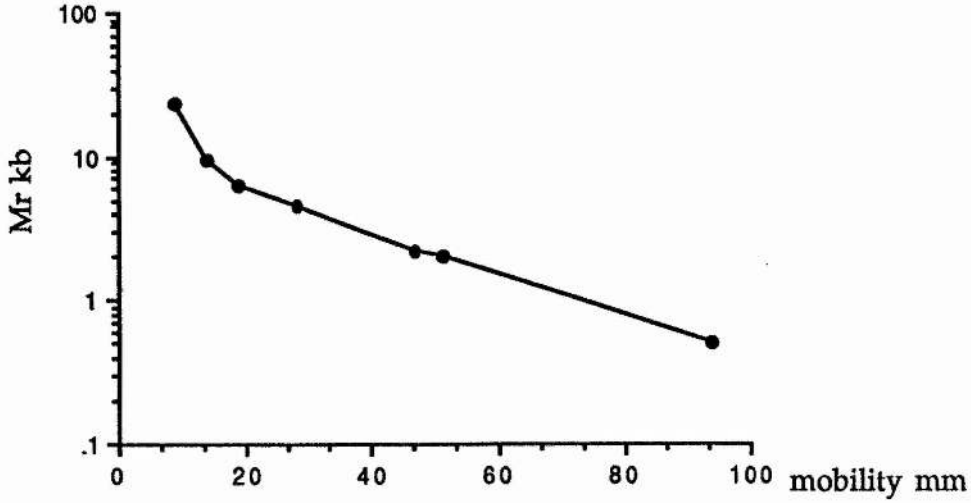
2. 5 .2 . b. Small DNA fragments (<600 bp)

DNA fragments of less than 600 bp were run on a 6% polyacrylamide gel buffered with 1 x TBE. Acrylamide solution (stored in the dark at 4° C) was 40% acrylamide, 2% bisacrylamide dissolved and filtered through two layers of Whatman No. 1 filter paper. The gel was poured between two glass plates separated with spacers which determine the thickness of the gel. Cross-linking occurs a short time after the catalysts N, N, N', N-Tetramethylethylenediamine (TEMED) and 25% Ammonium persulfate (APS) were added (75 µl and 132 µl were added for a 50 ml gel, respectively). The gel was pre-run in a vertical position with 1 x TBE as the running buffer until the current had equilibriated. DNA samples were loaded as previously with *Hae*III cleaved pUC DNA molecular weight markers. When electrophoresis was complete the gel was stained in 0.25 µg ml^{-1} EtBr for 20 min at room temperature. If necessary, destaining in distilled water was allowed to take place at 50° C until sharp bands of DNA could be seen.

2. 5. 2. c. RNA (Davis *et al*, 1986)

RNA was electrophoresed through a horizontal, formaldehyde denaturing 1.2% agarose gel to allow good size separation and resolution of single stranded RNA. For a 100 ml gel volume agarose was dissolved in 85 ml of DEP treated distilled water and 10 ml of 10 x MOPS buffer (200 mM MOPS pH 7.0, 50 mM NaOAc, 10 mM EDTA) with 0.25 µg

i)



ii)

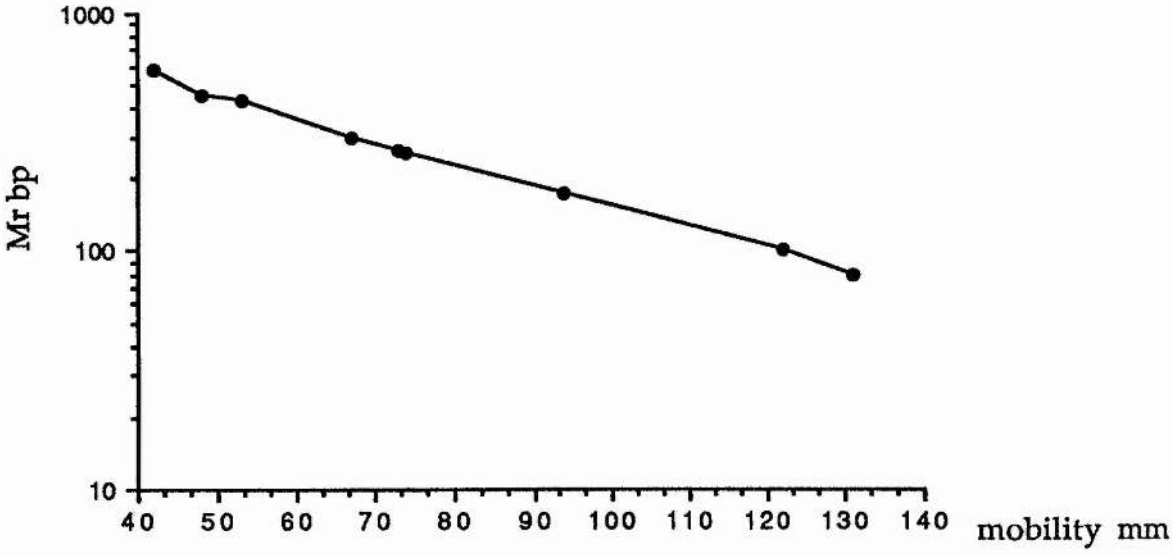


Figure 2.5 Graphs indicating the typical mobility of the molecular weight markers; i) *Hind*III cut λ DNA in 1.2% agarose gels and ii) *Hae*III cut pUC DNA in 2% agarose gels.

ml^{-1} of EtBr. Once the gel had cooled to 50° C, 5.4 ml of 40% formaldehyde was added. RNA samples were taken up in 5 volumes of sample buffer (0.76 ml formamide, 0.26 ml 40% formaldehyde, 0.18 ml DEP distilled water , 0.16 ml 10 x MOPS buffer) and 1 volume of loading buffer (2.5.2.a). The secondary structure of the RNA was broken by boiling the samples for 2 min, followed by immediate quenching on ice. Samples were loaded onto the gel as before (2.5.2.a) and generally run overnight at 15 V in 1 x MOPS buffer.

2. 5. 3. DNA extraction from low melting point gels (Sambrook *et al*, 1989)

A low melting point agarose gel (2.5.2.a) containing the DNA of interest was observed on a UV transilluminator and the DNA band required removed with a scalpel. The gel piece was cut into small fragments and melted in a 3 x volume of TE buffer in a 30 ml centrifuge tube at 70-75° C for 15 min. The agarose was removed by extraction with equal volumes of Tris saturated phenol until the interphase was clear. (Phases are separated by centrifugation at 4000 rpm for 5 min). Traces of phenol were removed by two extractions with equal volumes of chloroform containing 4% (w/v) IAA. Centrifugation at 4000 rpm was required for 2 min to separate the phases.

The volume of the DNA sample was reduced by mixing with an equal volume of butan-2-ol and discarding the upper phase. This was repeated until the volume was reduced to 1 ml, the sample being divided between two eppendorf tubes and each extracted with 1 ml of ether to remove all solvents. Heating the samples at 65° C for 15 min with the lids removed evaporated the remaining ether. The DNA was ethanol precipitated (2.2.4) and resuspended in a suitable volume of TE. Yields

were estimated by electrophoresis alongside a known quantity of DNA with similar molecular weight.

2. 5. 4. Transfer of nucleic acids to nylon membrane (Sambrook *et al*, 1989)

2. 5. 4. a Southern blotting

Southern blotting is a technique used to transfer DNA from an agarose gel to a nylon membrane placed directly above the gel. The DNA is denatured, neutralised and transferred in a high salt buffer by capillary action. DNA fragments larger than 15 kb require at least 18 h for transfer leading to dehydration of the gel and loss of transfer. To alleviate this, DNA is partially hydrolysed, by exposure to weak acid followed by strong base, prior to transfer. The resulting DNA fragments are approximately 1 kb.

The agarose gel containing the DNA to be blotted was photographed and inessential pieces removed if necessary. The gel was washed (shaking at room temperature) for two periods of 15 min in 0.25 M HCl in order to depurinate the DNA allowing hydrolysis of the phosphodiester backbone at these sites on exposure to alkali. The gel was agitated at room temperature for two periods of 15 min in denaturing solution (0.5 M NaOH, 1.5 M NaCl), required for separation of the two strands, enabling hybridisation to a radiolabelled probe. Finally, the gel was shaken in neutralising solution (0.5 M Tris HCl pH7.4, 3 M NaCl) for two washes each of 30 min. The gel was washed with distilled water after each buffer change.

Approximately 500 ml of 20 x SSC (3 M NaCl, 0.3 M sodium citrate pH 7.0) was poured into a flat bottomed tray. A small platform (the surface area of the gel or larger) was raised above the level of the 20 x SSC and two pieces of Whatman No. 1 filter paper, pre-soaked in 20 x SSC,

were layed across it, one extended length ways to come into contact with the 20 x SSC and the other extended width ways. The gel was placed upside down on the platform with the top right hand corner cut off for later identification. To prevent evaporation of the 20 x SSC, selophane was placed around the edge of the gel and over the sides of the tray forming an enclosed area around the gel. A piece of nylon membrane (Amersham International plc) the exact size of the gel, pre-soaked in 2 x SSC was placed on the surface of the gel ensuring the presence of no air bubbles. Two similar sizes of Whatman No. 1 paper were treated in a similar fashion, being placed individually over the nylon membrane. Several layers of dry paper towels were placed above the filter paper. Finally, a glass plate was placed on top of the paper towels with a 1 kg weight above it. Thus, a wick was set up with the 20 x SSC moving upwards carrying the DNA onto the nylon membrane.

The blot was left to stand overnight. After dismantling, the filter was allowed to air dry and the DNA cross linked by exposing the filter directly to UV light for 5 min (depending on the intensity of the UV).

2. 5. 4. b. Northern blotting

A formaldehyde denaturing gel containing the RNA to be blotted was photographed on a UV transilluminator as described previously (2.5.4.a). The RNA was denatured prior to, and during electrophoresis and therefore does not require treatment with alkali. Additionally, RNAs greater than 9 kb have been shown to transfer well and therefore do not require hydrolysis.

The gel was washed in 10 x SSC for two periods of 30 min to remove formaldehyde from the gel (formaldehyde may inhibit the transfer of the RNA to the nylon filter). The Northern blot was set up as for the Southern blot (2.5.4.a) and left standing overnight. After dismantling the

filter was treated in the same way as described for Southern blotting (2.5.4.a).

2. 5. 5. Hexaprime DNA labelling (Feinberg and Vogelstein, 1983)

Random sequence hexanucleotides are used to prime DNA synthesis catalysed by *E.coli* DNA polymerase (Klenow) in the presence of dNTPs and α-^{32}P dCTP. Template DNA was a single stranded, denatured restriction fragment with the sequence of interest.

Approximately 50 ng of the DNA to be labelled was boiled for 10 minutes in a total volume of 34 µl and subsequently plunged onto ice. To synthesize ^{32}P-labelled copies of both DNA strands, 10 µl of dCTP reaction mix (see below), 2 µl of 10 mg ml^{-1} BSA (NBL), 3 µl of α-^{32}P dCTP and 1 µl of 5 U µl^{-1} Klenow were added to the denatured DNA. The reaction proceeded for either 45 min at 37^{o} C or for at least 90 min at room temperature.

To separate the labelled DNA fragment from the unincorporated nucleotides the reaction mix was run through a Nick column (Pharmacia plc), comprising Sephadex G-50 (DNA grade). The Nick column was equilibriated with 1 ml of TE buffer before the DNA is added. The labelled DNA, moving through the column at a faster rate than unincorporated nucleotides, was collected at the bottom in a sterile eppendorf tube and extracted once with phenol : chloroform to remove impurities. The typical specific activity of each labelled probe was 1.5 x 10^{7} dpm µg^{-1} DNA. TE was continually added to the column until all the radioactivity was removed.

The labelled probe was denatured by boiling for 10 min before being quenched on ice and added to the hybridisation buffer (2.3.6.b).

The dCTP reaction mix was 100 µl of solution A (1 ml of 1.25 M Tris HCL pH 8.0, 0.125 M $MgCl_2$ with 18 µl of ß-mercaptoethanol, 5 µl each of

100 mM dGTP, dTTP and dATP in TE pH 8.0), 250 µl of solution B (2 M N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid] (HEPES) titrated to pH 6.6 with 4 M NaOH, 100 mM dNTPs in TE pH 8.0), and 150 µl of solution C (hexanucleotides (Pharmacia plc) at a concentration of 10 OD_{260} U ml^{-1} in TE pH 8.0).

2. 5. 6. Hybridisation to immobilised nucleic acids (Sambrook *et al*, 1989)

2. 5. 6. a. Northern hybridisation

Northern blots were pre-hybridised in hybridisation buffer for a period of 4-6 h to reduce non-specific probe hybridisation. The hybridisation temperature was usually 20-25° C below the melting temperature (Tm) of the probe. It has been shown that RNA : DNA hybrids have a Tm approximately 10° C higher than the equivalent DNA : DNA hybrid. Therefore it is particularly important, with Northern hybridisations, to use 40% deionised formamide (2.1.5.c) decreasing the Tm and providing less harsh conditions conditions at 42° C. High ionic strength solutions (5 x SSPE or 5 x SSC) maximise the rate of annealing of the probe. When formamide is included in the hybridisation 5 x SSPE is preferred due to its greater buffering capacity. Denhardt's reagent, 1% SDS and 250 µg ml^{-1} herring sperm DNA are used as agents to block the non-specific attachment of the probe to the surface of the filter.

The hybridisation solution for Northern blots consisted of 40% formamide, 5 x SSPE (3 M NaCl, 5 M $NaH_2PO_4.H_2O$, 20 mM EDTA, pH 7.4), 1% SDS, 5 x Denhardt's solution (50 x Denhardt's stock is 5 g Ficoll, 5 g poly vinyl pyrrolidone (PVP), 5 g BSA made up to 500 ml with distilled water and stored at -20° C) and 250 µg ml^{-1} of denatured herring sperm DNA. (Prepared by dissolving the solid sodium salt of the DNA to 10 mg ml^{-1}, sonicating and shearing through a 0.8 mm needle five times. It was

stored at -20° C and boiled for 10 minutes before adding to the hybridisation buffer)

For a filter of 12 x 11 cm, 20 ml of buffer was used, sealed in a plastic bag, excluding all air bubbles. Small volumes of hybridisation solution allow faster nucleic acid reassociation.

The conditions for hybridisation were identical. The labelled probe (2.5.5) was added to the bag after boiling and incubation proceeded for sufficient time to allow the probe to achieve 3 x $Co_t{}^1/_2$ as specified by the equation:

$$Co_t{}^1/_2 = {}^1/_X \times {}^Y/_5 \times {}^Z/_{10} \times 2$$

where X is the weight of the probe (μg), Y is the length of the probe (kb) and Z is the volume of the reaction (Sambrook *et al*, 1989)

In general, the washing conditions should be as stringent as possible. A combination of temperature and salt concentration determine the stringency. The washing temperature of Northern blots was 55° C with the salt concentration gradually decreased to 1 x SSC. 0.1% SDS was included in the washes. The stability of the RNA : DNA hybrid decreases as the salt concentration decreases. Only homologous hybrids should remain at the end of the washing procedure.

2. 5. 6. b. Southern hybridisation

Southern blots were pre-hybridised for a period of 4-6 h in a solution consisting of 5 x SSPE, 6% PEG 6000 Mr 7-9 K, 0.5% Marvel (Boots plc), 1% SDS, 0.1% $Na_2P_4O_7$ and 250 μg ml^{-1} of denatured herring sperm DNA prepared as before (2.5.6.b). Marvel replaces the Denhardt's as a blocking agent as it is generally easier to use and less expensive. PEG is added to

increase the rate of hybridisation by concentrating the probe. (Nucleic acids are excluded from the volume of the solution occupied by the polymer). Typically 20 ml of the solution was added to a 12 x 11 cm filter, which was pre-wetted with distilled water in a sandwich box. For a stringent or 'homologous' hybridisation, incubation was at 65° C. More relaxed conditions, for heterologous hybridisations, were obtained at 54° C.

After boiling, the labelled probe (2.5.5) was introduced directly into the pre-hybridisation buffer. Generally, hybridisation takes place overnight for at least the time specified by the equation in section 2.5.6.a.

Washing, initially at 5 x SSC and 0.1% SDS was decreased sequentially to 0.2 x SSC with 0.1% SDS for stringent hybridisations but was not carried any further if more relaxed conditions were required. The washing temperature was identical to the hybridisation temperature although it may be increased to remove background noise, if necessary.

2. 5. 7. Removing probes from membranes

Northern blots were stripped of their radioactivity and re-probed with radio-labelled *A.nidulans actA* gene fragments. The *actA* gene (Fidel *et al*, 1988), encoding an actin protein, is expressed by all the fungal strains under all growth conditions employed and therefore is a good internal control for monitoring equal transfer of RNA to the filter.

Filters were placed in 500 ml of distilled water with 0.1% SDS and brought to the boil. The process was repeated and filters checked with the mini-monitor for any remaining activity. Washes continued until no counts were detected and the filters stored in plastic bagging at -20° C until required.

2. 5. 8. Autoradiography

Both Northern and Southern blots were finally washed, at room temperature, in the SSC concentration of their last wash to remove residual SDS. Filters were sealed in plastic bags and autoradiographed against either Fuji film or Kodak XAR film, if a quick result was required. Films and filters were placed in light excluding cassettes between intensifying screens. Cassettes were stored at - 70° C until the film was developed in an automatic processor.

2. 5. 9. Strategy for screening the *A.nidulans* cDNA library

Initially, it was necessary to determine the concentration of the *A.nidulans* cDNA library, obtained from Dr. M. Innes (Cetus), in terms of the pfu ml^{-1}. Serial dilutions in TM buffer were made from the concentrated lysate, believed to be 10^8 $p.f.u.ml^{-1}$, and plated out using *E.coli* NM538 as the host cell by the following method.

A single colony of *E.coli* NM538 was grown in 10 ml of NZYDT media at 37° C over night and 100 µl of this culture used to inoculate a second 10 ml volume of NZYDT. These cells were grown at 37° C with 250 rpm orbital shaking until the OD_{600} was 0.45, pelleted at 3500 rpm and resuspended in 3 ml of TM buffer. A sterile 75 x 12 mm tube was used to mix 10 µl of the phage suspension with 150 µl of cells. The tubes were shaken gently (100 rpm) for 30 min at 37° C to encourage phage adsorbtion. After this period 3 mls of cool, molten, top agar (NZYDT with 0.7% agar) was added to each tube, the contents mixed and poured onto the surface of a dry plate. The plates were inverted and incubated at 37° C over night. The plaques formed were counted and from this the library was titred.

When screening a library for a relatively low copy number mRNA it is essential to ensure enough plaques to include a copy of each mRNA

transcribed, under the given set of conditions. From the Clarke and Carbon (1978) formula the number of plaques required was calculated:

$$N = \frac{\ln (1-P)}{\ln (1-f)}$$

where P is the desired probability of isolating the clone in question, f is the fraction of the genome in a single recombinant and N is the number of recombinants required to have P probability of obtaining the desired clone.

The average size of an *A.nidulans* mRNA was estimated as 3 kb and with a genome size of 2.6×10^7 bp this gave the value of f as $1.0^7 \times 10^{-4}$. Therefore 86,078 plaques were required to have a probability of 0.99 of locating the *crnA* cDNA.

The mRNA used as a template for the library was obtained from an *A.nidulans* culture grown at 37° C for 24 h. These were not thought to be optimal conditions for *crnA* gene expression (Brownlee and Arst, 1983) and therefore more than the calculated, 86,078 plaques were used to screen the library. It was decided to use ten, 9 cm petri dishes each with 1000 plaques. The library was diluted sufficiently to ensure the exact quantity of plaques would be obtained when plated out with *E.coli* NM538 as before.

The DNA peculiar to each plaque was transferred in duplicate to circular nylon filters with a 9 cm diameter (Amersham International plc). Each filter was labelled either 1 or 1' up to 10 or 10'. The first filter of each pair was marked with a dash in three assymetrical places around its edge. The second, primed filter was marked with a double dash in a similar

manner. The first filter was placed onto the surface of a plate, ensuring no creases or bubbles. While transfer was taking place the base of the plate was marked with the corresponding filter number and the edge of the plate marked where the dashes occur on the filter. After 3 min the filter was laid flat, plaque impressions uppermost, onto two layers of Whatman No. 1 soaked in denaturation buffer (2.5.4a) and left for 30 s. The filter was then treated with neutralising solution (2.5.4.a) for 2-3 min in the same way and washed in 3 x SSC for 15 min. After air drying, the filter was exposed to UV light for 5 min in order to crosslink the DNA.

The second filter of each pair was laid onto the surface of the corresponding plate ensuring that the double dashed lines on the filter did not coincide with the marks of the first filter. Transfer continued for 5 min and the filter position recorded on the sides of the plate as before. The duplicate filter was subsequently treated in the same way as the first.

Several nanograms of the probe DNA was placed onto the edge of one of the filters before it was treated with denaturing solution. Probe hybridising to this area indicates that the hybridisation is successful.

The filters were hybridised in 40 ml of buffer (2.5.6.b) at 60° C to the hexaprime radio-labelled (2.5.5) 3.3 kb *EcoRI/NruI* fragment of pSTA4 (Fig 2.2), and washed to 0.1 x SSC, 0.1% SDS (2.5.6.b). After autoradiography, ten positive signals were chosen for further analysis. An area containing 30-40 plaques around each signal (aligned using the marks on the sides of the plates corresponding to the dashes on the filters) was picked with a toothpick and placed into 1 ml of TM buffer in a sterile eppendorf tube.

In general, each plaque contains 10^7 pfu therefore it was necessary to dilute out the phage to a concentration that would result in a few hundred plaques after plating out. (The plaques corresponding to the area of one positive signal were confined to one plate). Once the correct

plating density was achieved the plaques were transferred to nylon membrane and treated as before.

After the secondary screen, positive signals were aligned with plaques as before. It was only necessary to pick a maximum of ten plaques to include every possibility of isolating the positive plaque. The required plating density of the tertiary screen was approximately 50 plaques per plate. This ensured alignment of the positive signal to a single plaque. A small scale liquid lysate was generated from each positive species and from this the bacteriophage DNA along with the insert was isolated (2.4.5).

Chapter 3

RESULTS

3. 1. 0. INTRODUCTION

The presentation of the results in this thesis are designed to follow the four sections of the "experimental programme" described in the INTRODUCTION (1.6.0).

First, the phenotype of the *A.nidulans crnA*1 mutant strain is examined (3.2.0). An outline of the structure of the *crnA* gene together with the transport protein it encodes is presented in section 3.3.0. Particular reference to some of the features discussed in the introduction are included, both at the nucleotide level (1.2.1) and AA motifs or domains within the polypeptide (1.5.2).

The regulation of *crnA* gene expression at the level of mRNA accumulation is examined in section 3.4.0 by studying mRNA levels on Northern blots. A method of preparing intact RNA was obtained (Cathala *et al*, 1983) favouring the use of this method. In parallel, nothing was known of the architecture of the *crnA* gene at this stage rendering the harnessing of *crnA* promoter sequences to the *lacZ* reporter gene difficult (1.6.2). The effect on *crnA* gene expression exerted by the regulatory genes, *nirA* and *areA* (3.4.3) and a brief examination of the regulatory role of the NR enzyme (3.4.4) is included in this section. In addition, the effect of altering the nitrogen and carbon source on *crnA* mRNA accumulation is determined.

An investigation of the regulation of *niiA* and *niaD* gene expression was conducted in a similar way to that of the *crnA* gene and detailed in section 3.5.0 To reinforce evidence for the autoregulatory role of NR a more extensive Northern blotting analysis is presented.

A comparison of the 5' non-coding sequences of the *niiA* and *niaD* genes of *A.nidulans* (3.6.1) is followed by a more detailed investigation

involving the *niiA-niaD* intergenic regions from *A.oryzae* and *A.niger* (3.6.2). The aim of the examination is the identification of putative *cis*-acting sites which may interact with regulatory gene products.

Finally, the results delivered in section 3.7.0 demonstrate that the *nit-4* gene of *N.crassa*, for induction of genes involved with nitrate assimilation, is able to complement a *A.nidulans nirA* loss of function mutation. Regulation of NR in the *nit-4* gene transformed strains is examined and compared with that of the wild type strain.

3. 2. 0. PHENOTYPE OF THE *A.NIDULANS crnA*1 STRAIN

A strain mutant at the *crnA* locus, *crnA*1, is resistant to chlorate and is able to utilise nitrate (1.4.2). Growth on nitrate and nitrite at concentrations of 0.1 mM, 1 mM, 2 mM and 5 mM was indistinguishable from the wild type (results not shown). In contrast, the addition of caesium chloride to the media at a concentration of 20 mM resulted in a significant loss of growth in comparison to the wild type with both nitrate and nitrite as sole nitrogen sources (Plate 3.1). The wild type demonstrated substantial loss of growth on caesium chloride concentrations greater than 50 mM although the effect on the growth of the *crnA*1 strain was exaggerated. It is considered here that caesium chloride may block nitrate and nitrite uptake by a second nitrate transport system (1.4.2) and with a mutation at the *crnA* locus the strain fails to grow on nitrate and nitrite. These results provide a defined phenotypic test for the *crnA*1 mutant strain and imply that the *crnA* gene may be responsible for nitrite transport in addition to nitrate.

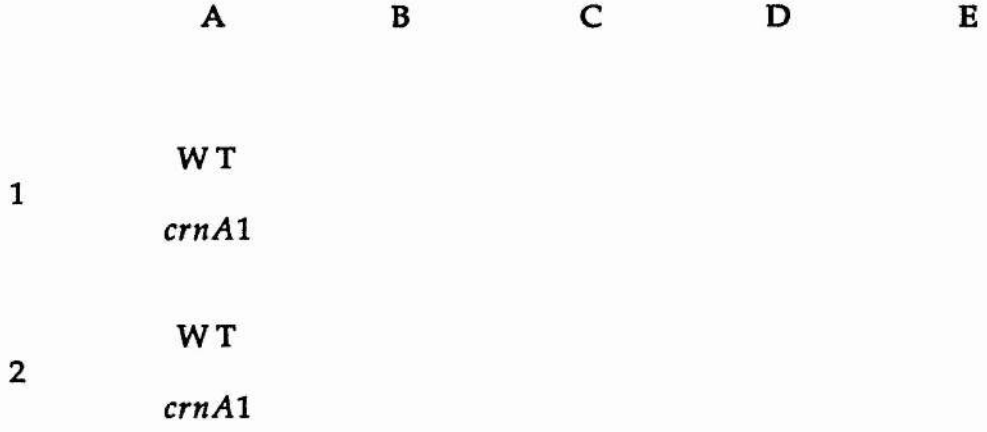
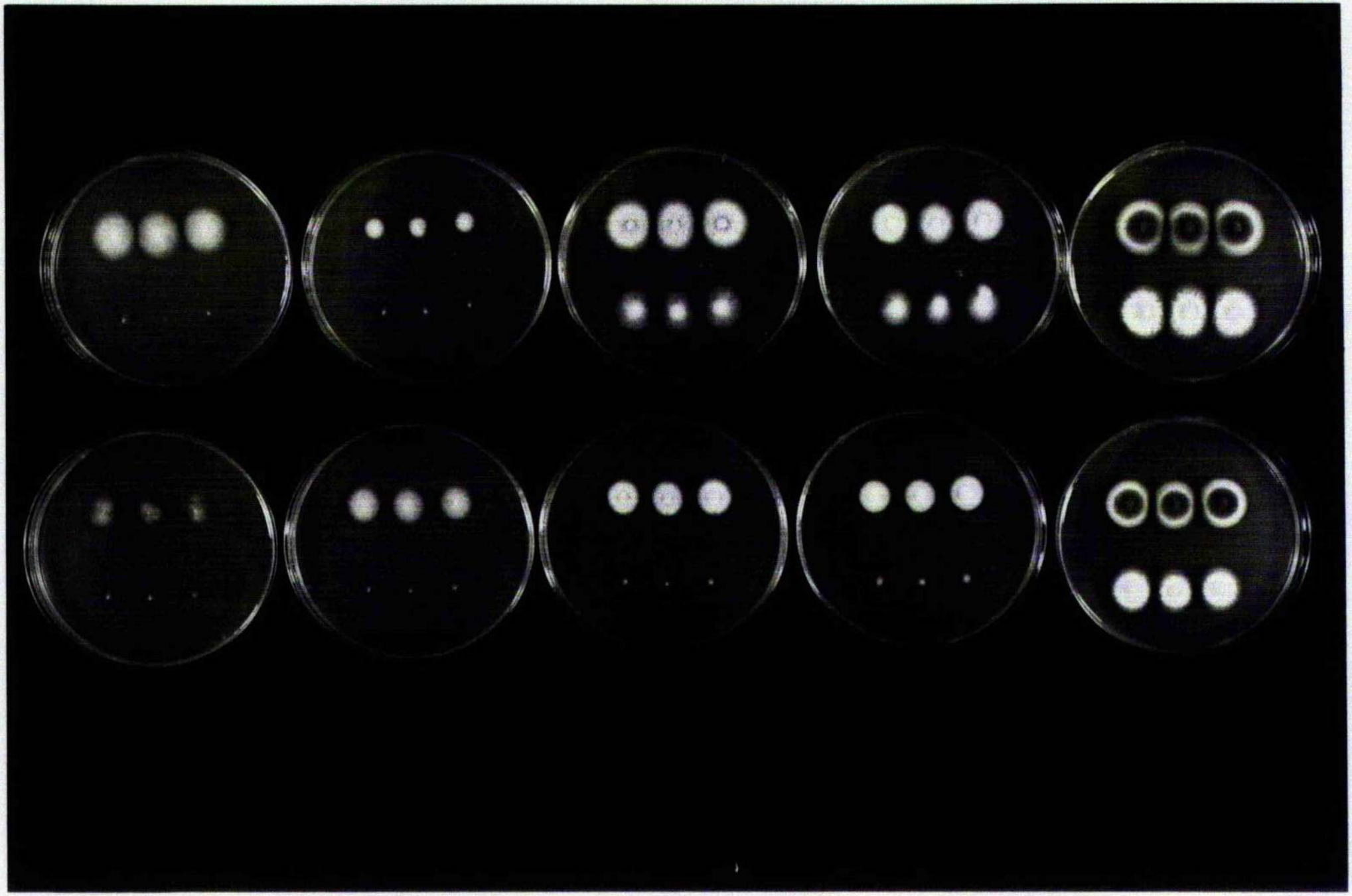


Plate 3.1 Growth characteristics of the wild type and *crnA*1 strains on nitrate and nitrite and varying concentrations of caesium chloride. From the key row 1 is MM with 5 mM nitrate and row 2 is MM with 5 mM nitrite. The caesium chloride concentrations are column A; 100 mM, B; 50 mM, C; 20 mM, D; 10 mM and E; 1 mM. The strains, wild type (top) and *crnA*1 (bottom) are shown in triplicate.



3. 3. 0. THE STRUCTURE OF THE *crnA* GENE AND PROTEIN

A 4.8 kb *EcoR*I fragment consisting of *crnA* gene-specific sequences had been cloned into the plasmid, pILJ16 (2.1.4.c), resulting in the construct pSTA4 (Fig 2.2). The adjacent *niiA* gene was known to extend to the *Nru*I site in pSTA4, inferring that the *crnA* gene is situated within the 3.3 kb *EcoR*I/*Nru*I (Fig 2.2) stretch of DNA (Johnstone *et al*, 1990).

3. 3. 1. Isolation of a *crnA* cDNA clone

A sequence analysis of the 3.3 kb *EcoR*I/*Nru*I pSTA4 fragment would identify the constitution of the *crnA* gene (3.3.2). In conjunction, and to aid in the elucidation, the *crnA* cDNA was isolated.

The *A.nidulans* cDNA library was synthesized in bacteriophage λgt10 and constructed using RNA from mycelia grown in nitrate for 24 h at 37° C (2.5.9). It seemed plausible these conditions were not ideal for the construction of a cDNA library with the aim of obtaining a copy of the *crnA* mRNA as there had been the suggestion that the *crnA* gene may be developmentally regulated (Brownlee and Arst, 1983) and possibly not expressed at 16 h (1.4.2). However, it was presumed that, taking into account the half-life of the *crnA* mRNA, there may be a small population of *crnA* transcripts remaining at 24 h.

From the Clarke and Carbon (1978) formula (2.5.9), the number of plaques required, N, for a 99 % probability of isolating a single *crnA* cDNA was calculated (2.5.9). The average size of an *A.nidulans* cDNA was presumed to be approximately 3 kb (taking into account the sizes of the *A.nidulans* transcripts previously isolated, e.g. *niiA* and *niaD* mRNAs (Johnstone *et al*, 1990)). This, as a fraction of the *A.nidulans* genome, 2.6 x 10^7 bp, provides a value for "f" with a corresponding value for "N" of 86,078. Therefore, to ensure optimal conditions for isolating at least one

copy of the *crnA* gene ten, 9 cm petri dishes were required, each with a density of 10,000 plaques.

The library was plated out by conventional methods (2.5.9), using *E.coli* NM538 as the host strain for bacteriophage λgt10 (2.1.4.c). Bacteriophage DNA was transferred, in duplicate, to nylon membrane and subsequently screened with the hexaprime radio-labelled (2.5.5) 3.3 kb *EcoR*I/*Nru*I fragment from pSTA4. The initial screen yielded 22 positive plaques. Areas of approximately 30-40 plaques surrounding ten of these were picked for the secondary screen with a lower plating density than the first (2.5.9). After the third and final screen, only one positive plaque remained. Loss of positive plaques may have occurred due to inaccurate alignment of the petri-dish with the positive signal on the X-ray film. A second plausible explanation is that the positive plaque may have died or lost the ability to infect and replicate during the time required for screening.

A lysate from the positive recombinant was used to prepare a large quantity of bacteriophage DNA (2.4.5). Digestion with *EcoR*I released the cDNA insert from the two bacteriophage "arms". The size of the fragment was measured as 1.35 kb by electrophoresis through an agarose gel alongside *Hind*III cut λ DNA as molecular weight markers (Fig 2.5). The cDNA of the *crnA* gene was expected to be approximately 1.6 kb, the size of the mRNA measured previously (Johnstone *et al*, 1990) by Northern blotting. Consequently, the 1.35 kb cDNA isolated from the library was deduced to be either a copy of a degraded mRNA or a copy of an entirely different gene with a similar sequence to the 3.3 kb *EcoR*I/*Nru*I fragment from pSTA4.

In an endeavour to discover which explanation was correct a restriction enzyme analysis was conducted. The DNA was refractory to digestion with the endonucleases *Pst*I, *Bgl*II and *Sma*I, however

restriction with *Sac*I resulted in the generation of a double DNA fragment. Such experimentation suggests that the *crnA* gene resides between the *Pst*I and *Eco*RI sites (Fig 2.2) but does not confirm that the 1.35 kb fragment is the *crnA* cDNA. In addition, this appears to contradict previous complementation analysis, which implied that the *crnA* gene, or at least the *cis*-acting regulatory sequences (1.2.1), are situated between the *Nru*I and *Sma*I sites (Greaves, 1990).

For ease of manipulation and propagation, the 1.35 kb putative *crnA* cDNA was cloned into pUC18 at the *Eco*RI site utilising competent

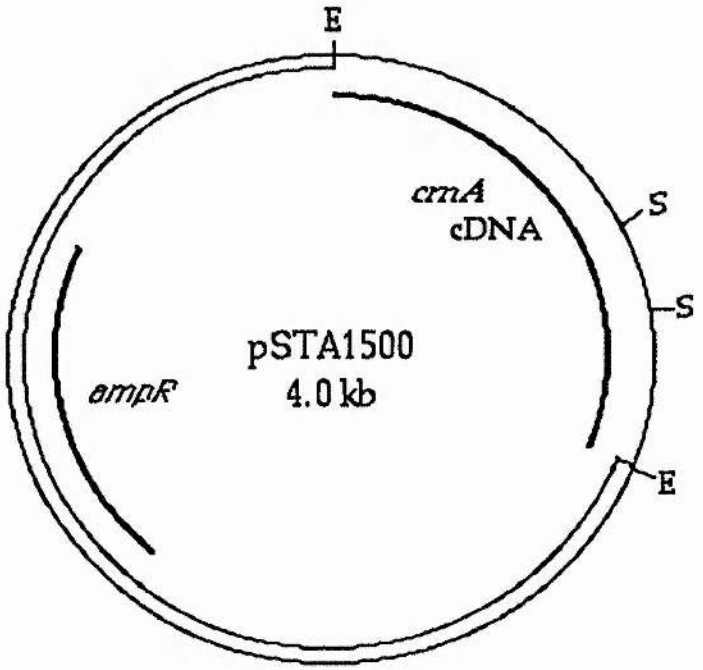


Figure 3.1 The vector pSTA1500 contains the *crnA* cDNA cloned into the plasmid, pUC18 at the *Eco*RI site (E). The position of the restriction site *Sal*I is indicated by S. The extent of the *ampR* gene and *crnA* cDNA are represented by a thick, black line. The concentric lines signify plasmid sequences.

Table 3.1 Comparative restriction fragment lengths of the *crnA* cDNA and genomic DNA.

restriction endonuclease(s)	fragment length in bp	
	genomic DNA	cDNA
*Hinc*II	975	-
	840	780
	645	430
	315 x2	125
	180	-
*Sph*I	1560	-
	1260	1000
	450	340
*Hinc*II/*Sph*I	975	-
	840	790
	450	340
	315 x2	125
	105	105
	90	-
*Sac*I/*Sal*I	1800	435
	915	670
	315	-
	255	255
*Hinc*II/*Sac*I	975	-
	645	420
	540	490
	315 x2	125
	255	255
	180	-

E.coli JM101 as the recipient cells (2.1.4.c). The resulting plasmid, termed pSTA1500 (Fig 3.1) was utilised in all subsequent *crnA* cDNA experimentation.

To further examine the cDNA a more rigorous restriction enzyme analysis was performed. The profile of fragments obtained with the cDNA, from various single and double digests, was compared with that of the recombinant DNA of pSTA4. In addition to a 1.2% agarose gel, the fragments were resolved on a 6% polyacrylamide gel (2.5.2.b) as many of the expected sizes were less than 600 bp and this would ensure a more accurate diagnosis. The molecular weight markers *Hind*III cut λ DNA (Fig 2.5) and *Hae*III cut pUC DNA were used where appropriate for the interpretation of fragment sizes. The results confirm the 1.35 kb fragment to be a copy of a degraded *crnA* mRNA and for ease of presentation are shown in table form (Table 3.1). The presence of one intron (IVS), within the *crnA* gene, of about 50 bp was identified by comparing the recombinant and cDNA restriction fragment lengths. In addition, the information obtained was used to assemble a restriction fragment map of the *Eco*RI/*Nru*I fragment, indicating the position of the cDNA (Fig 3.2).

3. 3. 2. DNA sequencing analysis

To discover the direction of transcription, identify the position of any further introns and more importantly perhaps, obtain the AA sequence of the CRNA protein to provide evidence for it being a membrane protein, it was necessary to sequence the complete *Eco*RI/*Nru*I fragment of pSTA4. Subclones of small (less than 1 kb) restriction fragments of the 3.3 kb *Eco*RI/*Nru*I fragment were ligated into M13mp18 and M13mp19 bacteriophage by Miss Carolyn Grieve. To allow complete sequencing of both DNA strands, single stranded DNA templates of the M13 clones were synthesized. The sequence was obtained and assembled

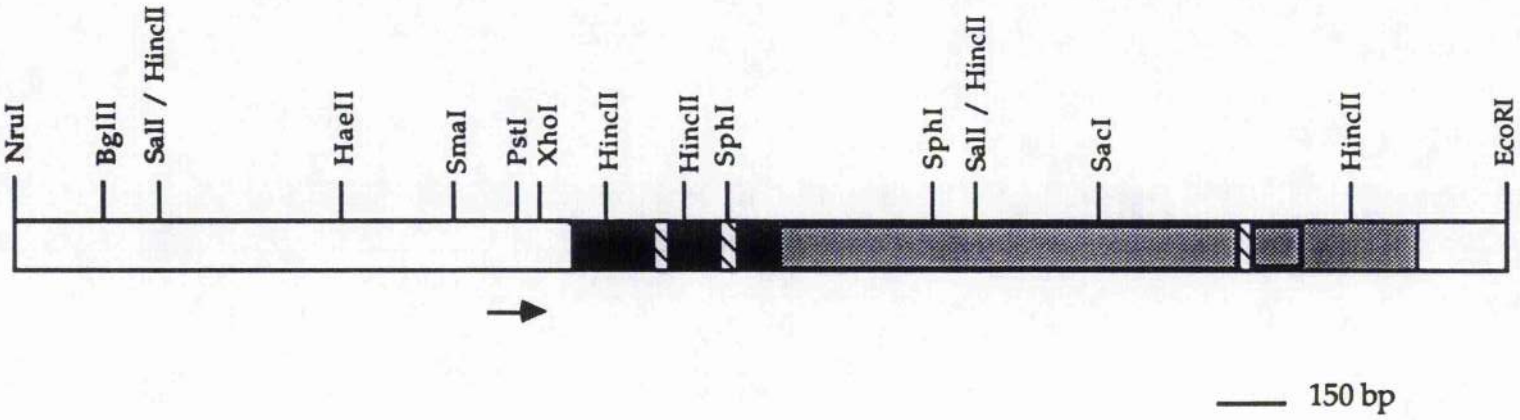
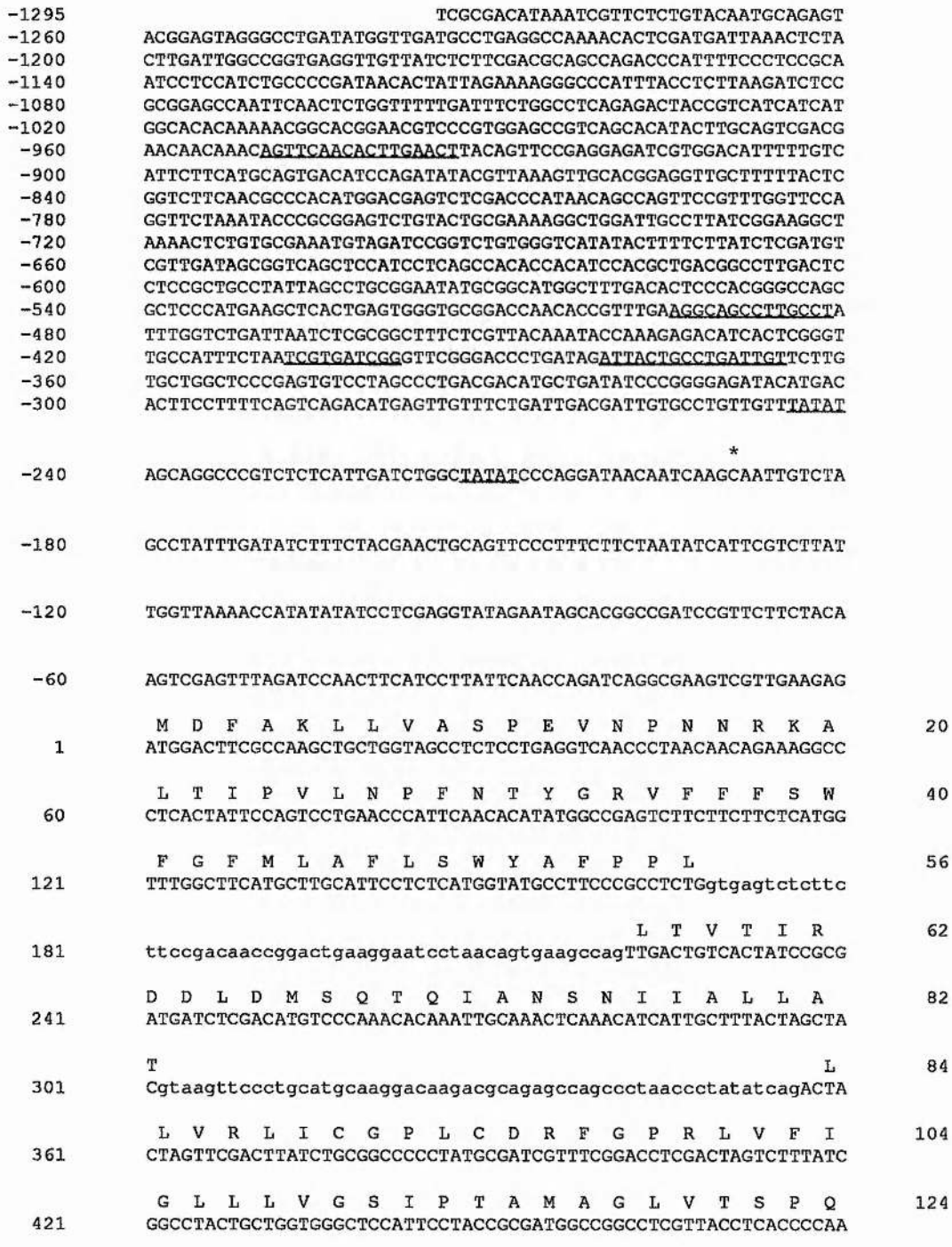


Figure 3.2 Restriction endonuclease map of the *crnA* gene and flanking regions. The transcriptional start site is represented with an arrow. The coding regions are enclosed in a black outline and the position of the introns are indicated with hatched boxes. The shaded area demonstrates the extent of the isolated cDNA.

by Dr. Shiela Unkles using the "Sequenase" kit (United States Biochemicals Co.) and shown in Fig. 3.3. The 2.0 kb gene, with three introns, encodes a polypeptide of 483 AAs.

Examination of the complete sequence shows the *crnA* gene to be transcribed from right to left, in the same direction as the *niiA* gene (Fig 3.2). An intron of 54 bp near the 3' end of the gene (Fig 3.3), corresponds with that identified previously (3.3.1). Sequence analysis of the cDNA fragment indicates the existence of an intron at this position. None of the motifs proposed to be necessary for polyadenylation (1.2.1) are evident although a long (30 nt) TG rich region, including just two cytosine residues, is situated 89 nt from the end of the cDNA (Fig 3.2). T-rich and TG-rich regions beyond the end of the coding sequences are associated with efficient poly(A) addition. A CATAAA motif, which has been demonstrated to allow polyadenylation at a reduced efficiency (Wickens, 1990), is present beyond the end of the cDNA. However, due to its position it is unlikely to play such a role. No consensus sequence, YGTGTTYY, for the termination of transcription could be identified; such a motif is rare in filamentous fungal genes (1.2.1).

A single open reading frame was identified, interrupted by either two or three introns. The termination codon is 215 bp upstream of the end of the cDNA. (Fig 3.3). That three, in frame, ATG codons are present, all resulting in a gene of sufficient length to satisfy the estimated size of the mRNA (3.4.1), provoked controversy. However, the favoured ATG at +1 (Fig 3.3), with a leader sequence in the region of 200 bp would transcribe an mRNA of 1.8 kb, corresponding more accurately to the results presented in this thesis (3.4.0). The remaining two ATG sites at +129 nt and +251 nt would require leader sequences of 230 bp and 120 bp respectively, before qualifying as candidates for a 1.6 kb message.


-1295 TCGCGACATAAATCGTTCTCTGTACAATGCAGAGT
-1260 ACGGAGTAGGGCCTGATATGGTTGATGCCTGAGGCCAAAACACTCGATGATTAAACTCTA
-1200 CTTGATTGGCCGGTGAGGTTGTTATCTCTTCGACGCAGCCAGACCCATTTTCCCTCCGCA
-1140 ATCCTCCATCTGCCCCGATAACACTATTAGAAAAGGGCCCATTTACCTCTTAAGATCTCC
-1080 GCGGAGCCAATTCAACTCTGGTTTTTGATTTCTGGCCTCAGAGACTACCGTCATCATCAT
-1020 GGCACACAAAAACGGCACGGAACGTCCCGTGGAGCCGTCAGCACATACTTGCAGTCGACG
-960 AACAACAAACAGTTCAACACTTGAACTTACAGTTCCGAGGAGATCGTGGACATTTTTGTC
-900 ATTCTTCATGCAGTGACATCCAGATATACGTTAAAGTTGCACGGAGGTTGCTTTTTACTC
-840 GGTCTTCAACGCCCACATGGACGAGTCTCGACCCATAACAGCCAGTTCCGTTTGGTTCCA
-780 GGTTCTAAATACCCGCGGAGTCTGTACTGCGAAAAGGCTGGATTGCCTTATCGGAAGGCT
-720 AAAACTCTGTGCGAAATGTAGATCCGGTCTGTGGGTCATATACTTTTCTTATCTCGATGT
-660 CGTTGATAGCGGTCAGCTCCATCCTCAGCCACACCACATCCACGCTGACGGCCTTGACTC
-600 CTCCGCTGCCTATTAGCCTGCGGAATATGCGGCATGGCTTTGACACTCCCACGGGCCAGC
-540 GCTCCCATGAAGCTCACTGAGTGGGTGCGGACCAACACCGTTTGAAGGCAGCCTTGCCTA
-480 TTTGGTCTGATTAATCTCGCGGCTTTCTCGTTACAAATACCAAAGAGACATCACTCGGGT
-420 TGCCATTTCTAATCGTGATCGGGTTCGGGACCCTGATAGATTACTGCCTGATTGTTCTTG
-360 TGCTGGCTCCCGAGTGTCCTAGCCCTGACGACATGCTGATATCCCGGGGAGATACATGAC
-300 ACTTCCTTTTCAGTCAGACATGAGTTGTTTCTGATTGACGATTGTGCCTGTTGTTTATAT
*
-240 AGCAGGCCCGTCTCTCATTGATCTGGCTATATCCCAGGATAACAATCAAGCAATTGTCTA
-180 GCCTATTTGATATCTTTCTACGAACTGCAGTTCCCTTTCTTCTAATATCATTCGTCTTAT
-120 TGGTTAAAACCATATATATCCTCGAGGTATAGAATAGCACGGCCGATCCGTTCTTCTACA
-60 AGTCGAGTTTAGATCCAACTTCATCCTTATTCAACCAGATCAGGCGAAGTCGTTGAAGAG
M D F A K L L V A S P E V N P N N R K A 20
1 ATGGACTTCGCCAAGCTGCTGGTAGCCTCTCCTGAGGTCAACCCTAACAACAGAAAGGCC
L T I P V L N P F N T Y G R V F F F S W 40
60 CTCACTATTCCAGTCCTGAACCCATTCAACACATATGGCCGAGTCTTCTTCTTCTCATGG
F G F M L A F L S W Y A F P P L 56
121 TTTGGCTTCATGCTTGCATTCCTCTCATGGTATGCCTTCCCGCCTCTGgtgagtctcttc
L T V T I R 62
181 ttccgacaaccggactgaaggaatcctaacagtgaagccagTTGACTGTCACTATCCGCG
D D L D M S Q T Q I A N S N I I A L L A 82
241 ATGATCTCGACATGTCCCAAACACAAATTGCAAACTCAAACATCATTGCTTTACTAGCTA
T L 84
301 CgtaagttccctgcatgcaaggacaagacgcagagccagccctaaccctatatcagACTA
L V R L I C G P L C D R F G P R L V F I 104
361 CTAGTTCGACTTATCTGCGGCCCCCTATGCGATCGTTTCGGACCTCGACTAGTCTTTATC
G L L L V G S I P T A M A G L V T S P Q 124
421 GGCCTACTGCTGGTGGGCTCCATTCCTACCGCGATGGCCGGCCTCGTTACCTCACCCCAA

```
                G  L  I  A  L  R  F  F  I  G  I  L  G  G  T  F  V  P  C  Q      144
481     GGACTGATTGCCCTGCGCTTCTTCATCGGCATCCTCGGCGGCACATTCGTTCCCTGCCAA

                V  W  C  T  G  F  F  D  K  S  I  V  G  T  A  N  S  L  A  A      164
541     GTCTGGTGCACAGGGTTTTTTGACAAGAGTATAGTTGGGACAGCCAACTCCCTAGCTGCC

                G  L  G  N  A  G  G  G  I  T  Y  F  V  M  P  A  I  F  D  S      184
601     GGTCTAGGTAACGCTGGTGGCGGTATCACATACTTCGTCATGCCGGCCATCTTCGACTCC

                L  I  R  D  Q  G  L  P  A  H  K  A  W  R  V  A  Y  I  V  P      204
661     CTCATCCGTGACCAAGGCCTCCCCGCACACAAGGCCTGGCGCGTCGCCTACATCGTCCCC

                F  I  L  I  V  A  A  A  L  G  M  L  F  T  C  D  D  T  P  T      224
721     TTTATCTTAATCGTTGCCGCCGCCCTAGGCATGCTCTTCACTTGCGATGACACCCCGACT

                G  K  W  S  E  R  H  I  W  M  K  E  D  T  Q  T  A  S  K  G      244
781     GGAAAATGGTCCGAGCGGCACATCTGGATGAAGGAGGATACCCAGACAGCATCTAAAGGC

                N  I  V  D  L  S  S  G  A  Q  S  S  R  P  S  G  P  P  S  I      264
841     AACATTGTCGACCTTAGCTCTGGTGCACAGTCCTCCCGTCCCTCCGGACCCCCTTCCATT

                I  A  Y  A  I  P  D  V  E  K  K  G  T  E  T  P  L  E  P  Q      284
901     ATTGCGTACGCCATTCCCGACGTCGAAAAGAAAGGCACCGAGACTCCACTAGAACCTCAG

                S  Q  A  I  G  Q  F  D  A  F  R  A  N  A  V  A  S  P  S  R      304
961     TCCCAGGCAATCGGCCAATTCGACGCCTTCCGCGCAAACGCCGTTGCCTCTCCCTCCCGC

                K  E  A  F  N  V  I  F  S  L  A  T  M  A  V  A  V  P  Y  A      324
1021    AAGGAGGCTTTTAATGTTATATTCAGCCTCGCAACGATGGCCGTTGCAGTCCCCTACGCC

                C  S  F  G  S  E  L  A  I  N  S  I  L  G  D  Y  Y  D  K  N      344
1081    TGCTCCTTTGGGTCTGAGCTCGCAATCAACTCGATCCTGGGCGACTATTATGACAAGAAC

                F  P  Y  M  G  Q  T  Q  T  G  K  W  A  A  M  F  G  F  L  N      364
1141    TTCCCGTACATGGGCCAAACGCAGACCGGCAAGTGGGCCGCTATGTTCGGGTTCCTTAAT

                I  V  C  R  P  A  G  G  F  L  A  D  F  L  Y  R  K  T  N  T      384
1201    ATTGTCTGTCGTCCGGCAGGTGGATTTCTGGCGGATTTCCTTTACCGGAAAACGAACACG

                P  W  A  K  K  L  L  L  S  F  L  G  V  V  M  G  A  F  M  I      404
1261    CCCTGGGCTAAGAAACTCCTCCTCTCGTTTTTAGGTGTTGTCATGGGTGCATTCATGATT

                A  M  G  F  S  D  P  K  S  E  A  T  M  F  G  L  T  A  G  L      424
1321    GCAATGGGTTTCTCAGATCCAAAGTCCGAAGCGACTATGTTTGGTCTTACTGCCGGGTTG

                A  F  F  L  E  S  C  N  G  A  I  F  S  L  V  P  H  V  H  P      444
1381    GCCTTCTTTCTTGAGTCTTGCAATGGGGCAATATTTTCGCTTGTACCACATGTTCATCCT

                Y  A  N  G                                                      448
1441    TATGCTAATGGTgtgtttttcccactttgcgctcaactttgaaaggacattcaactaaca

                      G  S  S  P  A  W  W  V  D  S  G  T  S  A  V  S  S  S      466
1501    gttacagGGATCGTCTCCGGCATGGTGGGTGGATTCGGGAACCTCGGCGGTATCATCTTC
```

```
          P  S  S  S  A  I  V  I  T  T  T  R  A  A  S  G  F                483
1561      GCCATCATCTTCCGCTATAGTCATCACGACTACGCGCGCGGCATCTGGATTCTAGGTGTT

1621      ATTTCCATGGCCGTGTTTATCTCTGTCTCCTGGGTTCGGCCTGTGCCGAAAAGTCAGATG

1681      AGGGAGTAGTTGACCGTGTGATTGGGTTCTCAAGAGGCGTTTGGGTTCGGTTTTGTTCTG

1741      TTTTGTTTAGCGATATGATACCATGATATACTATAGACTTGAGTGATGATTGTGAATGCA

                *
1801      TTTTGGCCATACCTGATGTCCACCTTCATACCCGTACTCTCTTTTCTATTATGATATAGT
1861      TGGCCGTTGGAGGAACTGCATAAATTCCAGGCTTAACTACCCACTCCTTCGGGCTAATAC
1821      TCCCCGAACTGAGGGGCAACATCTAGGTCTGAGGTAACATACATCTTCGCCCTCCCTTGG
1881      ACCTTCACTCAAAATATCTTCAAGCCAGTTTACAAAAAAAACTTTCAGTGATCCCCGGGAA
2041      TTC
```

Figure 3.3 Nucleotide and amino acid sequences of the *crnA* gene. Motifs referred to in the text are underlined. The stars indicate the sites of transcriptional initiation and termination. The nucleotides are numbered on the left and the AAs on the right.

Although leader sequences of this length in filamentous fungal genes are unusual, they are not unknown (1.3.2) and further experimentation was necessary before the correct transcriptional start site could be elucidated.

In contrast to nuclease S1 analysis the method of primer extension was found easier to manipulate. Using this method Dr. Shiela Unkles identified the transcriptional start site as the cytosine residue at position -190 bp (Fig 3.3). The first ATG downstream of the transcriptional start site is generally the site of translational initiation, inferring the ATG at +1 (Fig 3.3) represents the start of the coding sequence and the open reading frame is interrupted by three introns (Unkles *et al*, 1991). The sequence surrounding the ATG of GAGATGGA does not correspond to the consensus necessary for targeting the ribosome to the correct site. The only conserved nucleotide is the guanine residue immediately 3' of the ATG.

The determination of the 5' end of the *crnA* gene allowed an examination of the complete *crnA* nucleotide and AA sequences. The length of the *crnA* gene can be inferred as 2.0 kb with an mRNA of 1.84 kb, consistent with measurements from Northern blots in this study (3.4.0). Two introns are located towards the 5' end of the coding region (Fig 3.3), amounting to a total of three introns within the *crnA* gene. The introns are small: 53 bp, 55 bp and 55 bp for IVS1, IVS2 and IVS3 respectively (numbered 1 - 3, in order from the 5' to the 3' ends of the gene), in common with other *A.nidulans* genes (1.2.1). Each was recognised by the presence of the consensus sequence for lariat formation (1.2.1) near the 3' end of each intron: IVS1-TCCTAAC, IVS2-CCCTAAC, and IVS3-AACTAAC. The two nucleotides at the 5' (GT) and 3' (AG) extremities of the three introns are identical:

Intron	5'splice site	3' splice site
IVS1	TG'gt----------	----------ag'TT
IVS2	AC'gt----------	----------ag'AC
IVS3	GT'gt----------	----------ag'GG

The nucleotide sequence upstream of the transcriptional initiation site of the *crnA* gene was examined for core promoter sequences and potential *cis*-acting sequences or protein-binding sites; hairpin loops, repeating sequences and areas of similarity with the corresponding regions of the *niiA* and *niaD* genes (3.6.1), using the 'analyseq' program (author Rodger Staden) on the University of St.Andrews vax. The *crnA* promoter would appear to consist of one or possibly two TATA boxes (1.2.1) situated 19 bp and 51 bp upstream from the transcriptional start site. In contrast to many other filamentous fungal genes, a TC rich region (1.2.1) is located 40 bp downstream of the transcriptional start site. Neither a CAAT box nor a GC box would appear to be present in the promoter sequences although this not surprising as their functional significance in lower eukaryotes is unclear (1.2.1).

Two short sequences 292 nt and 734 nt upstream of the transcriptional start site were identified as having the potential to form a hairpin loop. These are AGGCAGCCTTGCCT and AGTTCAACACTTGAACT respectively. Neither sequence is present upstream of the *niiA* or *niaD* genes reducing the probability that their role is functional. Many different repeating sequences of six nucleotides in length occur throughout the length of the *crnA* gene and upstream regions (too many to demonstrate clearly), making it difficult to identify any as having a role in the regulation of *crnA* gene expression. The sequence TATCTA, a related form, or its complement occurs at least twice in areas upstream of the *niaD* gene and the *nit-3* gene of *N.crassa* that

have been demonstrated to bind the product of the *N.crassa nit-2* regulatory gene (Fu and Marzluf, 1990). The 5' non-coding sequences of the *crnA* gene include several copies of the consensus TATCTA. In particular, one area 5' to the transcriptional start site contains one copy of TTTCTA and also the complementary sequence TAGATT. This area is located between -378 and -415 nt.

The *niiA* and *niaD* 5' flanking sequences were compared and shown to share eight regions of similarity (3.6.1). It is possible that these motifs represent *cis*-acting sequences, binding sites for the products of the regulatory genes *nirA* and *areA*. The *crnA* gene is shown to be regulated in a similar manner to the *niiA* and *niaD* genes (3.4.0 and 3.5.0) and therefore it was proposed that the eight motifs may also be common to

Table 3.2 Nucleotide sequence comparison of the *niiA*, *niaD* and *crnA* 5' non-coding regions. The distance of the motif from the translational start site is measured from the 3' end and does not include the ATG.

Motif	Position from translational start site		Sequence	Identity	Length(nt)
4	*niiA*	-169	GTCATTGGCCCATTT	53%	15
	crnA	-379	ATTACTGCCTGATTG		
4	*niaD*	-187	GTTTTCGTCTCATTT	53%	15
	crnA	-379	ATTACTGCCTGATTG		
5	*niiA*	-198	TCGCTGATTCG	73%	11
	crnA	-405	TCG-TGATCGG		
5	*niaD*	-216	TCT-TGATTTG	73%	11
	crnA	-405	TCG-TGATCGG		

the 5' non-coding sequences of the *crnA* gene. A computer-aided comparison revealed little similarity between the eight motifs and the sequences upstream of the *crnA* gene. However, a sequence of 10 bases, TCGTGATCGG, at -215 nt from the *crnA* transcriptional start site demonstrates 73% identity with motif 5 of the *niiA* and *niaD* genes (Fig 3.3 and Table 3.2). Additionally, a second sequence, ATTACTGCCTGATTGT, at -189 nt displays 53% similarity with motif 4 of both the *niiA* and *niaD* genes (Fig 3.3 and Table 3.2).

3. 3. 3. Codon usage

The AA coding region was analysed for significant codon bias

Table 3.3 The *crnA* gene codon usage. Amino acids are represented as the single letter code. An asterisk indicates a stop codon.

F	TTT	12.	S	TCT	10.	Y	TAT	5.	C	TGT	1.
F	TTC	26.	S	TCC	14.	Y	TAC	6.	C	TGC	7.
L	TTA	3.	S	TCA	7.	*	TAA	0.	*	TGA	0.
L	TTG	2.	S	TCG	7.	*	TAG	1.	W	TGG	10.
L	CTT	8.	P	CCT	8.	H	CAT	2.	R	CGT	4.
L	CTC	13.	P	CCC	11.	H	CAC	2.	R	CGC	6.
L	CTA	10.	P	CCA	6.	Q	CAA	7.	R	CGA	3.
L	CTG	10.	P	CCG	6.	Q	CAG	5.	R	CGG	2.
I	ATT	11.	T	ACT	9.	N	AAT	4.	S	AGT	1.
I	ATC	17.	T	ACC	7.	N	AAC	14.	S	AGC	2.
I	ATA	4.	T	ACA	8.	K	AAA	5.	R	AGA	2.
M	ATG	14.	T	ACG	6.	K	AAG	11.	R	AGG	0.
V	GTT	10.	A	GCT	9.	D	GAT	8.	G	GGT	11.
V	GTC	15.	A	GCC	23.	D	GAC	11.	G	GGC	18.
V	GTA	3.	A	GCA	16.	E	GAA	3.	G	GAA	8.
V	GTG	2.	A	GCG	6.	E	GAG	7.	G	GGG	6.

(Table 3.3). All sense codons are used with the exception of AGG, for arginine. There is a slight bias (65%) towards a pyrimidine at the third position, when a choice is permitted, in general agreement with other filamentous fungal gene codon usage (1.2.1). However, as a whole, there is little tendency for excessive utilisation of one codon, where more than one codes for an AA. This would support the results of section 3.3.0 indicating that the *crnA* gene is not highly expressed.

3. 3. 4. Protein analysis

The nucleotide sequence of the *crnA* gene was translated to the AA sequence using the 'analyseq' computer program (3.3.3). The 483 AA protein has a molecular weight of 51,769 Da (Table 3.4). The large percentage of apolar residues, namely 46%, and their grouped arrangement (Fig 3.3) provides evidence that the *crnA* gene encodes a membrane protein, supporting the view that it is a nitrate transporter.

To examine the similarity of the *crnA* protein with other transport proteins a number of "best fit" AA alignments were performed using the UWGCG program (Wisconsin University) on the Daresbury vax. No substantial AA identity could be found with the *E.coli* melibiose (Yazyu *et al*, 1984), lactose (Buchel *et al*, 1980) and proline (Nakao *et al*, 1987) transport proteins nor, surprisingly, the nitrate transporter of *Synechococcus* PCC 7942 (Omata, 1990). The glucose (Celenza *et al*, 1988), galactose (Nehlin *et al*, 1989), arginine (Hoffman, 1985), histidine (Tanaka and Fink, 1985) allantoate (Rai *et al*, 1988), uracil (Jund *et al*, 1988) and purine cytosine (Weber *et al*, 1990) transport proteins of *S.cerevisiae* and the maltose permease of *Saccharomyces carlsbergensis* (Yao *et al*, 1989) also demonstrated little similarity. In addition, the phosphate transporter of *N.crassa* (Mann *et al*, 1989) and the proline (Sophianopoulou and Scazzocchio, 1989) and quinic acid (Hawkins *et al*, 1988) transporters of

A.nidulans indicated no more than 38% identity with the *crnA* AA sequence. Neither the human glucose transporter (Mueckler *et al*, 1985) nor the transport protein in which mutations lead to the disease cystic fibrosis (Riordan *et al*, 1989) showed any extensive similarity with the CRNA protein. Surprisingly, perhaps, the lactose carrier of *E.coli* proved to have the greatest similarity, 41%, with the *crnA* encoded polypeptide, although it is questionable whether this figure is significant.

Table 3.4 Amino acid usage of the *crnA* encoded polypeptide. The number of times each amino acid occurs in the polypeptide (n) and its percentage of the whole protein (n%) are indicated.

		483 AA	Mr: 51769 Da	
			n	n%
A	ala	alanine	54	11.2
L	leu	leucine	46	9.5
G	gly	glycine	43	8.9
S	ser	serine	41	8.5
F	phe	phenylalanine	38	7.9
I	ile	isoleucine	32	6.6
P	pro	proline	31	6.4
T	thr	threonine	30	6.2
V	val	valine	30	6.2
D	asp	aspartic acid	19	3.9
N	asn	aspargine	18	3.7
K	lys	lysine	16	3.3
R	arg	arginine	16	3.3
M	met	methionine	14	2.9
Q	gln	glutamine	12	2.5
Y	tyr	tyrosine	11	2.3
E	glu	glutamic acid	10	2.1
W	trp	tryptophan	10	2.1
C	cys	cysteine	8	1.7
H	his	histidine	4	0.8

It was expected that a eukaryotic anion transporter, for example the *N.crassa pho-4* (Mann *et al*, 1990) encoded protein, may share extensive similarity with the CRNA protein, since proteins transporting similar ions would be expected to have a similar structure. That no sequence comparison revealed significant identity implies that many transport proteins have evolved separately and the similarity seen here is due solely to the high percentage of hydrophobic AA residues in each protein.

It was necessary to classify the protein depending on the number of membrane spanning domains (1.5.2). This was conducted according to the method of Eisenberg *et al*, 1984 in which windows of 21 AAs are examined for the highest hydrophobicity. Ten such regions, consisting mainly of apolar AAs, were identified, with hydrophobicity values as shown (Table 3.5), inferring that this protein is polytopic (1.5.2). To support these data, a hydrophobicity plot of the CRNA protein was calculated. The ten hydrophobic membrane-spanning domains, corresponding to those already identified, are shown (Fig 3.5). Further, these membrane spanning domains can themselves be classified as either globular, surface seeking or integral by comparing their hydrophobicity with their hydrophobic moment (1.5.2), providing an insight to the structure of the protein within the membrane (Table 3.4). The putative model of the *crnA* encoded polypeptide shows ten groups of 21 AAs thought to form α-helices crossing the membrane, connected by loops of relatively polar residues in the cytoplasm and on the external surface (Fig 3.5). The orientation of the protein within the membrane was deduced according to the theory that positively charged AAs generally remain in the cytoplasm (1.5.3). In eukaryotes, only the charge distribution between the N-terminus and loop1 (connecting hydrophobic regions 1 and 2) is important for determining protein orientation (Dalbey, 1990). The N-

Table 3.5 Classification of each of the hydrophobic membrane-spanning domains and the charge carried by each of the connecting loops. The hydrophobicity, H and the hydrophobic moment, μH for each of the membrane spanning domains of the *crnA* encoded polypeptide, as determined by Eisenberg *et al* (1984) are indicated. A plot of μH against H determines the classification of the domains surface, S or integral, I. In addition, the charge carried by each of the connecting loops and termini is indicated.

Domain	H	μH	Classification	Charge
N-terminus				+2
Hydrophobic 1	0.72	0.32	S/I	
Loop 1				-2
Hydrophobic 2	0.52	0.20	I	
Loop 2				+1
Hydrophobic 3	0.75	0.08	I	
Loop 3				+1
Hydrophobic 4	0.69	0.05	I	
Loop 4				0
Hydrophobic 5	0.64	0.17	I	
Loop 5				+1
Hydrophobic 6	0.82	0.20	I	
Loop 6				-4
Hydrophobic 7	0.59	0.06	I	
Loop 7				-1
Hydrophobic 8	0.52	0.4	S/I	
Loop 8				+4
Hydrophobic 9	0.70	0.10	I	
Loop 9				-1
Hydrophobic 10	0.57	0.13	I	
C-terminus				0

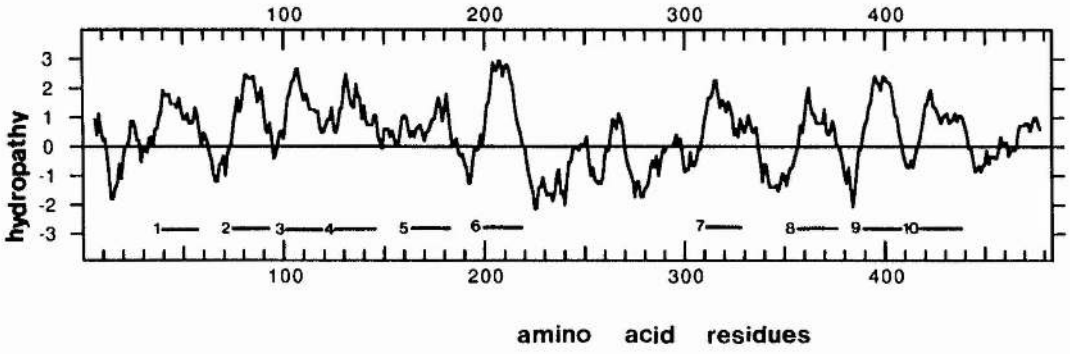


Figure 3.4 Hydropathy plot of the *crnA* encoded polypetide. Hydropathy values for a window of 13 AAs were calculated by the method of Kyte and Doolittle (1982). The solid lines indicate the ten regions most likely to form a membrane spanning helix. Hydrophobic regions correspond to the positive index numbers.

terminus of the *crnA* encoded protein has a greater positive charge than loop 1, +2 compared with -2 respectively (Table 3.5), suggesting that both termini and even numbered loops are not transported across the membrane, instead remaining in the cytoplasm.

It was envisaged that the proteins encoded by the *niaD* and *crnA* genes may share a common AA sequence necessary for the interaction with a nitrate molecule. No evidence of any such motif could be found from a "best fit" AA alignment.

Examination of the *crnA* AA sequence reveals that no ATP-binding motif (1.5.2) is present. A possible interpretation is that nitrate transport is not directly coupled to ATP utilisation. However, it is possible that other motifs may be responsible for binding ATP, as suggested by an interpretation of the sequence data of the "cystic fibrosis" transport protein (Riordan *et al*, 1989) and further motifs may remain to be identified.

It is considered that the *crnA* encoded polypeptide may be assembled at the plasma membrane and inserted either post-translationally or co-translationally. This conclusion is as a result of there being no potential glycosylation motifs (1.5.2) within the AA sequence. The possibility that the protein encoded by the *crnA* gene is assembled on the ER and transported to the plasma membrane by the Golgi is therefore remote, since most proteins translocated by this route are glycosylated.

According to the model assumed by the *crnA* encoded polypeptide (Fig 3.5) seven proline residues are distributed between the 10 membrane-spanning domains. Two are situated in domain 1 with the remainder occurring uniquely in each of domains 3, 4, 5, 6 and 8. It is possible that these residues along with glycine, serine and arginine result in the disruption of the α-helices within the membrane, thereby serving a role in transport (1.5.2). Certainly, the secondary structure of the CRNA

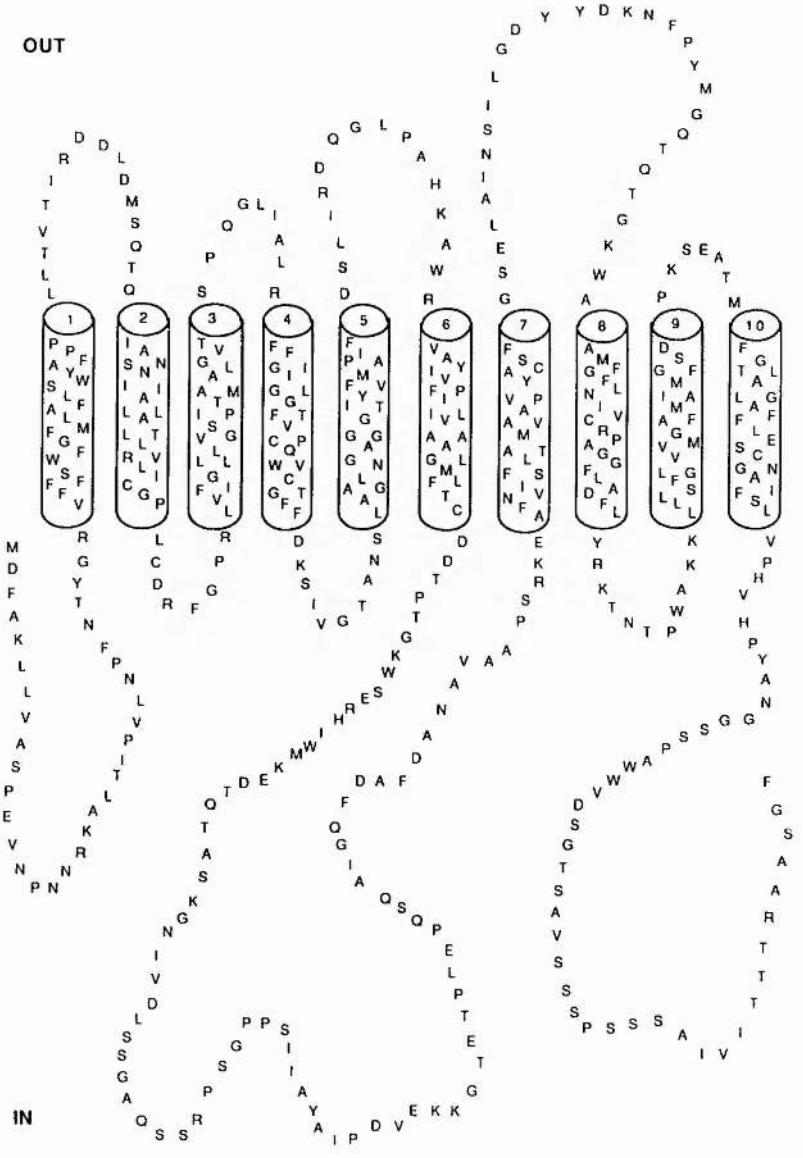


Figure 3.5 Model for the secondary structure of the *crnA* encoded polypeptide, calculated from the method of Eisenberg *et al* (1984). The ten putative membrane-spanning helices are represented as cylinders.

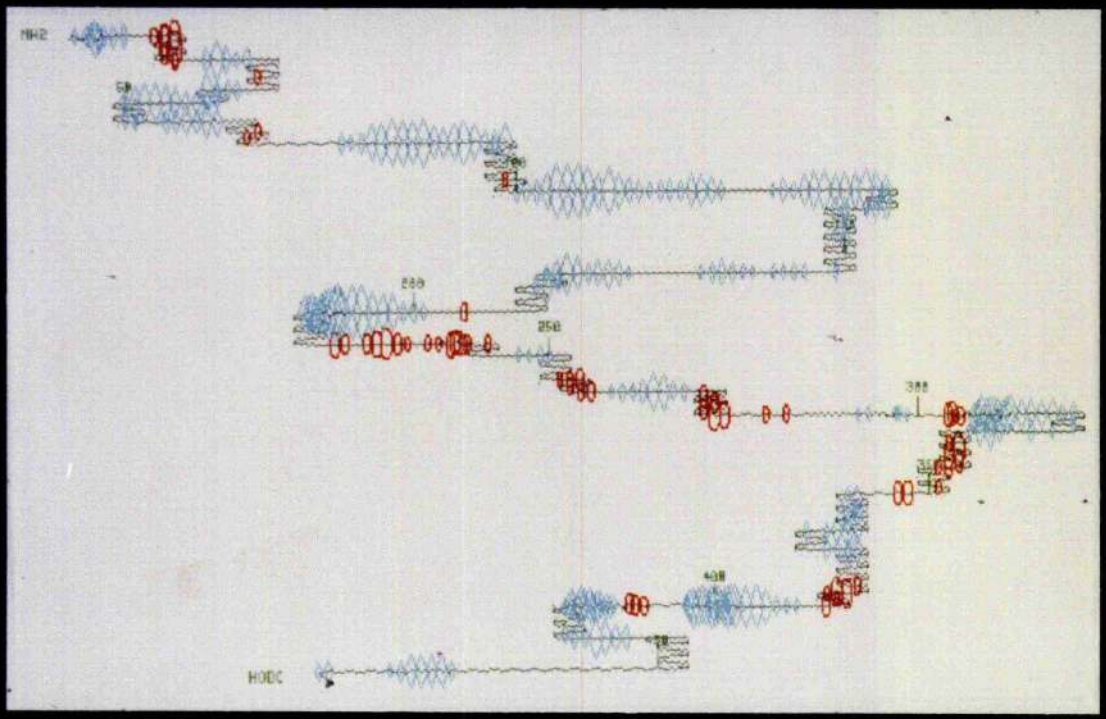
Plate 3.2 Computerised prediction of the secondary structure of the *crnA* encoded polypeptide. The straight lines represent ß-sheets and the position of α-helices are indicated by a zig-zag line. The blue diamonds are representative of the hydrophobic domains. In contrast, red circles are indicative of particularly hydrophilic areas.

protein, as predicted by the rules of Garnier, Osguthorpe and Robson (1978) (UWGCG program), does not clearly show the ten membrane-spanning α-helices (Plate 3.2). According to this model the secondary structure of the *crnA* encoded polypeptide is composed mainly of ß-sheet structures. This is a reasonable interpretation due to the relative abundance of the AAs valine, isoleucine, phenlyalanine and tryptophan, all favouring the formation of ß-sheets. Further support for this model is obtained from the observation that many glycine residues reside in the transmembrane domains. As explained, these may disrupt the α-helical structure. However, the computer model does not predict the structure the protein may assume when inserted into a membrane and further work is necessary before a true prediction can be made (4.7.3).

3. 4. 0. REGULATION OF *crnA* GENE EXPRESSION

Regulation of gene expression at the level of mRNA accumulation may be studied by examining mRNA levels on Northern blots (1.3.1) The expression of the *crnA* gene was found to be regulated at this level allowing an investigation into the mechanism of regulation.

The *crnA* cDNA hybridises to two messages, one of 1.1 kb and the other of 1.8 kb, measured using *Hind*III cut λ DNA molecular weight markers. A quantity of 20 μg of total RNA isolated from nitrate grown mycelia is required for the detection of both messages (Fig 3.6). A reduced amount leads to the complete loss of the smaller message although the 1.8 kb mRNA may be detected with 5 μg of RNA. The 1.8 kb message, shown to have greater abundance than the 1.1 kb mRNA, is considered to be that of the *crnA* mRNA as its size compares well with that predicted from the sequencing data (3.3.2). However, the smaller transcript is of unknown origin.


NH2
50
200
250
300
400
HOOC

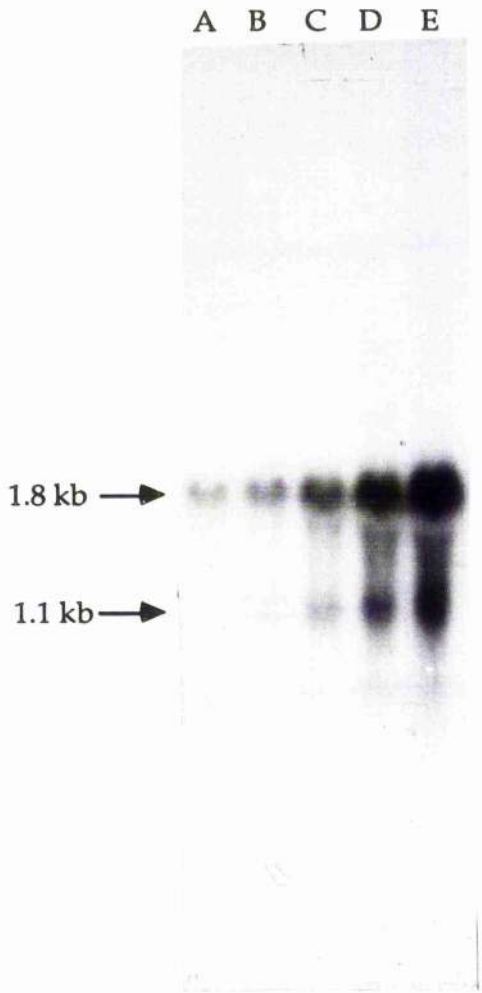


Figure 3.6 Northern blot of *A.nidulans* total RNA isolated from wild type grown in MM with 10 mM nitrate. Lane A; 5 µg, lane B; 10 µg, lane C; 20 µg, lane D; 50 µg, lane E; 100 µg. RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled *crnA* cDNA fragment from pSTA1500.

An initial investigation into the two messages was conducted by increasing the post-hybridisation washing temperature from 55° C to 65° C in 0.1 x SSC and 0.1% SDS. Both transcripts disappeared at the same rate suggesting that the nucleotide sequence of the two messages is equally similar to the probe DNA (data not shown). To examine the possibility that the smaller message may be from a second nitrate transport gene (1.4.2) Northern and Southern blots of *crnA* deletion mutants (Fig 1.6) were hybridised against the *crnA* cDNA. Total genomic *A.nidulans* DNA for Southern blotting was digested with endonuclease *Sal*I and electrophoresed through a 0.8% agarose gel. The hybridisation profile (Fig 3.7) of the wild type, lane A and *crnA-niiA-niaD* Δ507 strain, lane B, is identical to that obtained by previous workers for the *crnA* gene (Johnstone *et al*, 1990). The larger band of 3.8 kb, fragment S1 (Fig 1.6) is unaltered in the mutant, in contrast to the 1.8 kb fragment, S2 (Fig 1.6) showing an increase in size. That no other band is detected implies that the *A.nidulans* genome does not include two nitrate transport genes of similar sequence. In support of this, the Northern blot (Fig 3.7) of total RNA from nitrate grown mycelia of the strains; wild type, lane A, Δ507, lane B, and Δ506 (Fig 1.6), lane C shows that, in contrast to the wild type, both the 1.8 kb and 1.1 kb transcripts are absent in the deletion mutants. The Northern blot was stripped of radioactivity and re-probed with the *A.nidulans* actin gene (Fig 3.7 iii) demonstrating equal loading and transfer of RNA. The actin gene is expressed in all these strains and is therefore a good internal control (2.5.7). These results not only demonstrate that the Δ507 deletion extends into the upstream regulatory sequences of the *crnA* gene, i.e. into S2 (Fig 1.6), but also imply that the 1.1 kb message is either transcribed from within the region of the *crnA* gene or it is an artefact (4.2.0). It is clear from the sequencing data that only one open reading frame, that of the *crnA* gene itself, exists on either strand

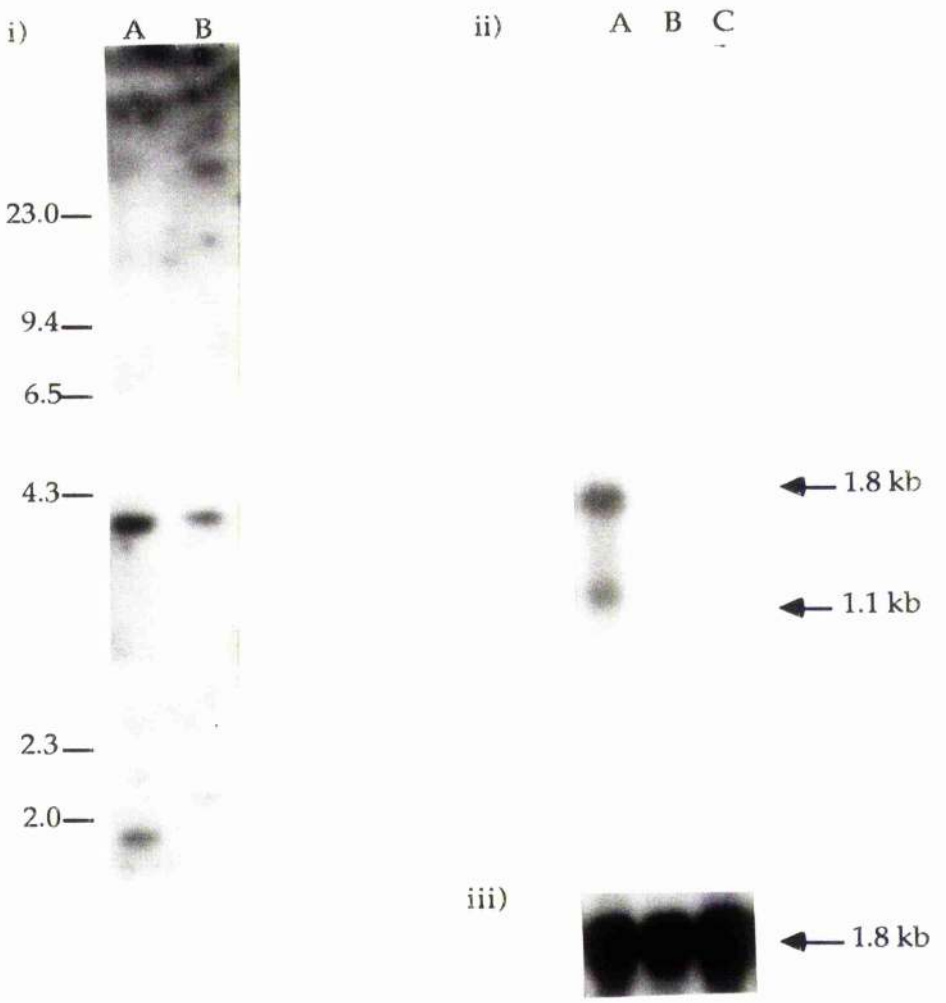


Figure 3.7 i) Southern blot of *A.nidulans* DNA isolated from the strains, lane A; wild type and lane B; Δ507. 20 µg of DNA was digested to completion with *Sal* I, electrophoresed on a 0.8% agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled cDNA fragment from pSTA1500.

ii) and iii) Northern blot of *A. nidulans* total RNA isolated from the strains, lane A; wild type, lane B; Δ506, lane C; Δ507. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, ii) *crnA* cDNA fragment from pSTA1500 and iii)*A.nidulans actA* fragments.

between the *EcoR*I and *Nru*I sites (Fig 3.3) of pSTA4. Therefore, it is plausible to assume that the 1.1 kb transcript is derived from the *crnA* gene itself.

3. 4. 1. The effect of different nitrogen sources on *crnA* gene expression

The induction and repression of the *crnA* gene by different nitrogen sources is shown (Fig 3.8 i). The 1.8 kb full length *crnA* transcript is regulated, in contrast to the 1.1 kb message which would appear to be constitutively expressed and therefore omitted from the remainder of the results, but discussed further in section 4.2.0.

Nitrate is responsible for the induction of the *crnA* gene, as a message is detected under these conditions (Fig 3.8 i lane A). In contrast, ammonium, glutamate or proline as the nitrogen source do not induce the expression of the *crnA* gene (Fig 3.8 i lanes B, C and D, respectively). To test the theory that the presence of ammonium ions result in nitrogen metabolite repression (1.4.7.b) mycelia were grown in media containing nitrate, together with proline, as nitrogen sources and also in nitrate together with ammonium (2.1.3.a). A *crnA* message is observed under the former conditions, lane E (Fig 3.8 i), but is absent when ammonium is introduced to the media, lane F (Fig 3.8 i). This supports the view that ammonium mediates repression of functions involved with nitrate assimilation (1.4.7.b) and that its effects override the ability of nitrate to convey induction. In contrast, proline shows the characteristics of a neutral nitrogen source: neither induction nor repression.

The Northern blot was stripped of radioactivity and re-probed with the actin gene (Fig 3.8 ii) demonstrating that RNA transfer in lane A was less than the remaining lanes. Induction of *crnA* gene expression by nitrate therefore is greater, relatively, than indicated in Fig 3.8 i lane A.

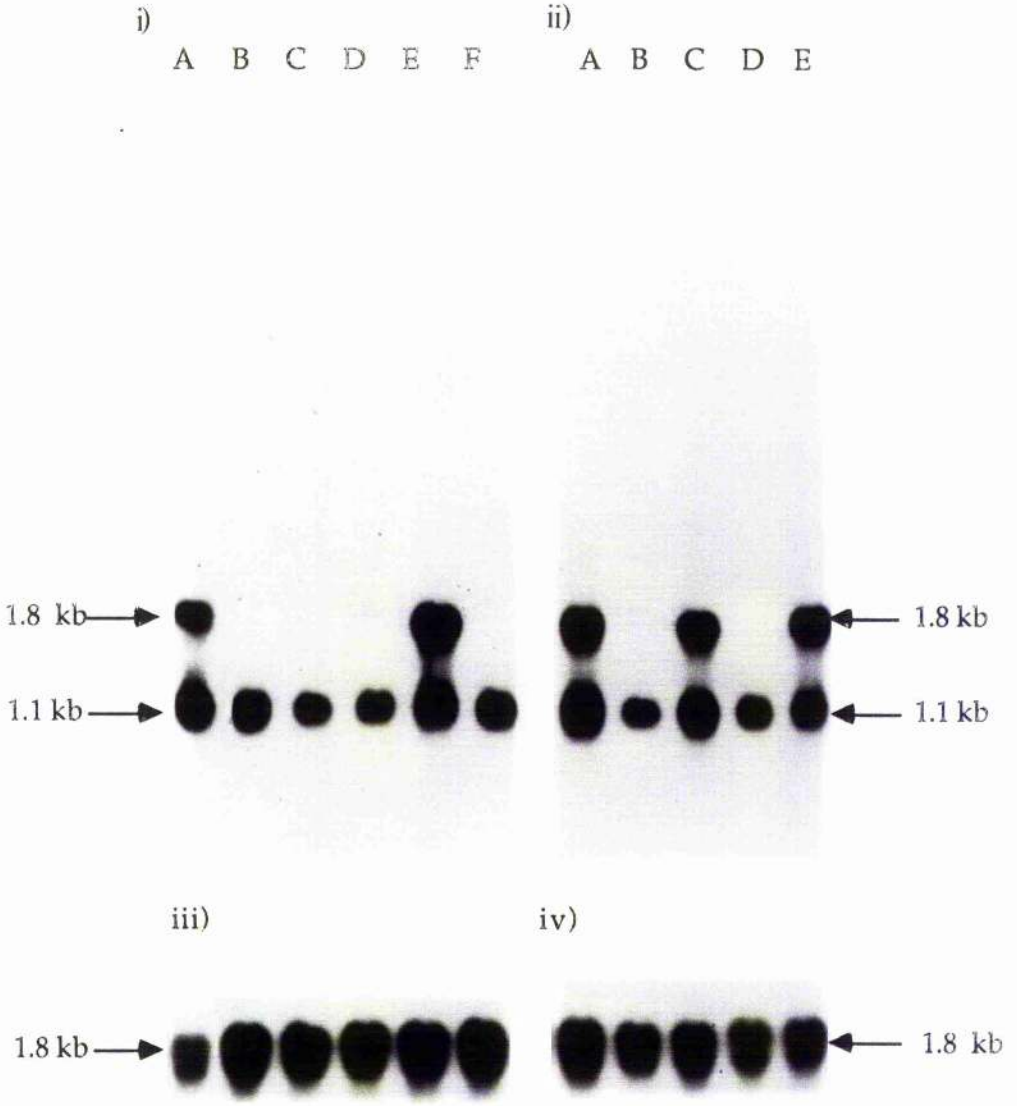


Figure 3.8 Northern blots of *A.nidulans* total RNA isolated from various strains. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) and iii) *crnA* cDNA fragment from pSTA1500 and ii) and iv) *A.nidulans actA* fragments.

i) and iii) wild type grown in MM with, lane A; 10 mM nitrate, lane B; 10 mM ammonium, lane C; 10 mM glutamate, lane D; 10 mM proline, lane E; 10 mM nitrate with 10 mM proline, lane F; 10 mM nitrate with 10 mM ammonium.

ii) and iv) lane A; wild type grown in MM with 10 mM ammonium and transferred to 10 mM nitrate, lane B; *nirA*1 grown under similar conditions to the wild type, lane C; *nirA*^{c} grown in MM with 10 mM glutamate, lane D; *areA*19 grown under similar conditions to the wild type, and lane E; *xprD* 1 grown in MM with 10 mM nitrate and 10 mM ammonium.

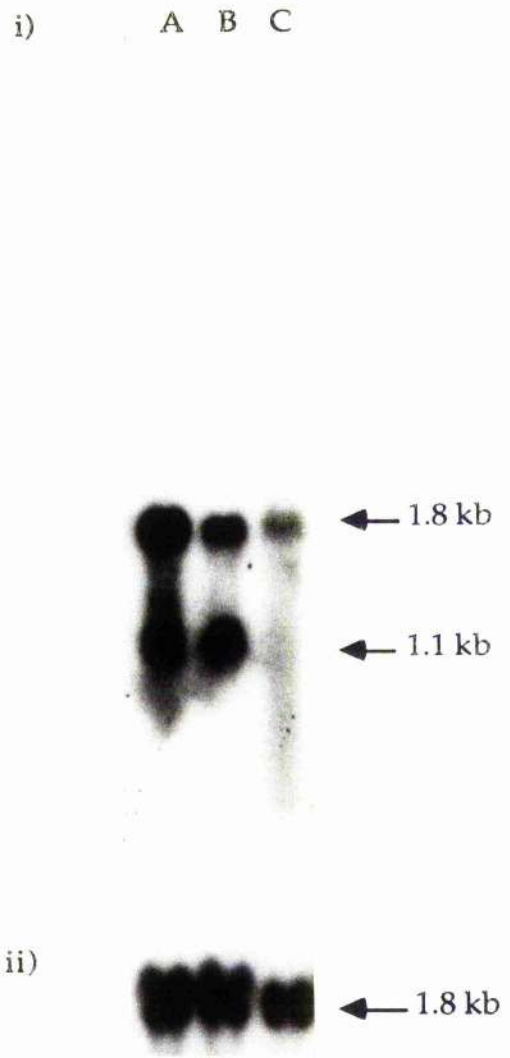


Figure 3.9 Northern blot of *A.nidulans* total RNA isolated from the wild type grown in MM with; lane A; 10 mM nitrate, lane B; 10 mM nitrite and lane C; 10 mM ammonium and transferred to 10 mM chlorate. 10 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *crnA* cDNA fragment from pSTA1500 and ii) the *A.nidulans actA* fragments.

Interestingly, Fig 3.9 demonstrates that nitrite (lane B) the reduced form of nitrate is able to induce the expression of the *crnA* gene. It is possible that the *crnA* gene acts as a specific transport system for nitrate and nitrite. Additionally, that a *crnA* transcript is observed when chlorate is supplied (lane C) would suggest that the mechanism by which nitrate is recognised is not specific and molecules of a similar shape and size may be mistaken for nitrate. However, that nitrite and chlorate appear to induce the expression of the *crnA* gene, does not necessarily signify that the *crnA* gene product is capable of their transport, i.e. they may be gratuitous inducers. It is worth noting that the 1.1 kb transcript is not detected in RNA extracted from mycelia grown in an environment with 10 mM chlorate. This result is unexpected as the 1.8 kb transcript is generally under tighter regulation (3.4.1 and 3.4.3). One explanation for this is that the quantity of RNA loaded in lane C was not great enough to allow the detection of the 1.1 kb mRNA, as described previously (3.4.0). The *actA* gene probe hybridises less strongly in lane C (Fig 3.9 ii) supporting the above possibility and suggesting that the apparent induction of the *crnA* gene by chlorate is greater, relatively, than shown in Fig 3.9 i.

3. 4. 2. The effect of glucose on *crnA* gene expression

Hynes (1973) observed that glucose as a carbon source was required for the synthesis of the enzyme NR. To investigate the extension of this theory to the nitrate transporter, mycelia were transferred from MM with nitrate, to a similar environment in which the 1% glucose had been replaced by 20 mM glutamate as the carbon source (2.1.3.b). Harvesting took place after t = 0, 15 and 30 min and 1, 2 and 3 h. RNA extractions and Northern blots were performed as before (3.4.1).

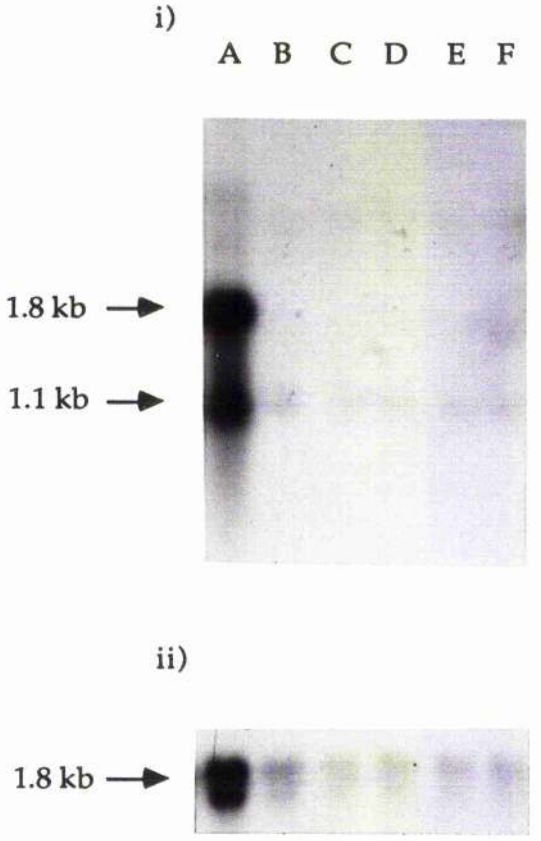


Figure 3.10 Northern blot of *A.nidulans* wild type total RNA isolated from mycelia grown in MM with 10 mM nitrate and transferred to similar media with 20 mM glutamate replacing glucose as the carbon source for time periods of, lane A; 0 minutes, lane B; 15 minutes, lane C; 30 minutes, lane D; 1 hour, lane E; 2 hours, lane F; 3 hours. 20 µg of RNA was electrophoresed on a 1.2 % formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *crnA* cDNA fragment from pSTA1500 and, ii) *A.nidulans actA* fragments.

The results (Fig 3.10 i) indicate a very rapid loss of message synthesis from time t = 0 (lane A) to time t = 15 min (lane B), sustained until time t = 3 h (lane F). That no RNA is detected at t = 15 min suggests rapid degradation of the remaining mRNA in the cytoplasm. This implies that the half life of the *crnA* mRNA is less than 7.5 min and supports the results of section 3.4.1, suggesting tight regulation.

A significant decrease in the quantity of actin mRNA (Fig 3.10 ii) from time t = 0 (lane A) to time t = 15 min (lane B) is also seen and it follows that a considerable number of genes are likely to require glucose for expression. This argues against any particular significance for glucose on the regulation of the genes of the nitrate cluster.

3. 4. 3. Regulation of *crnA* gene expression by the *nirA* and *areA* genes

The *nirA* and *areA* genes both encode positive acting regulatory proteins (1.4.7). It is known the products of the two genes, in addition to the co-effectors NR and glutamine, are necessary for the induction and repression of the nitrate assimilatory structural genes, viz *niiA* and *niaD*. A mutation within the *areA* or *nirA* genes leading to pleiotropic loss of function forfeits expression of the structural genes despite the presence of the inducer.

The loss of function mutants, *nirA*1 and *areA*19, were grown on MM and 10 mM ammonium (2.1.3.a) then shifted to 10 mM nitrate for 4 hours. Fig 3.8 ii shows that neither of the strains, *nirA*1 (lane B) or *areA*19 (lane D) are capable of *crnA* gene expression under inducing conditions. Conversely, the wild type grown under similar conditions, does show *crnA* gene expression (lane A). These results suggest that both the *nirA* and *areA* genes are required for *crnA* gene expression. This differs from a previous proposal in which the expression of the *crnA*

gene was said to be independent of the *nirA* gene (Scazzocchio and Arst, 1989).

In support of these findings a mutation within the *nirA* gene leading to the constitutive synthesis of NR, namely *nirA*^{c}1, was grown on glutamate. The *crnA* gene is found to be constitutively expressed in this mutant background (lane C), unlike the wild type (Fig 3.8 i, lane C), demonstrating unequovically that a functional *nirA* gene is required for the expression of the *crnA* gene.

In addition, a strain with an *xprD*1 mutation, known to be a 3' truncation of the *areA* gene (1.4.7.b), resulting in the derepression of NR activity, was grown on nitrate and ammonium (2.1.3.a). The expression of the *crnA* gene is shown to be derepressed in this mutant background providing evidence for the requirement of a functional *areA* gene in *crnA* gene expression. To verify these results it was confirmed that induction and repression of *crnA* gene expression in the *nirA*^{c}1 and *xprD*1 mutant backgrounds by nitrate and ammonium, as sole nitrogen sources, was normal, i.e. wild type (results not shown).

That a functional *nirA* gene was not capable of *crnA* gene expression in the strain with a mutation at the *areA*19 locus indicates that both functional *nirA* and *areA* genes are required simultaneously for *crnA* gene expression.

Equal loading and transfer of RNA in each lane (Fig 3.8 iii) was determined by stripping the filter of radioactivity and re-probing with the *actA* gene fragments (Fig 3.8 iv).

3. 4. 4. Regulation of *crnA* gene expression by NR

It has been shown that several *niaD* mutants are resistant to chlorate and bring about the constitutive synthesis of a functional NiR and a *niaD* encoded polypeptide (Cove, 1976 a and b). This led to the

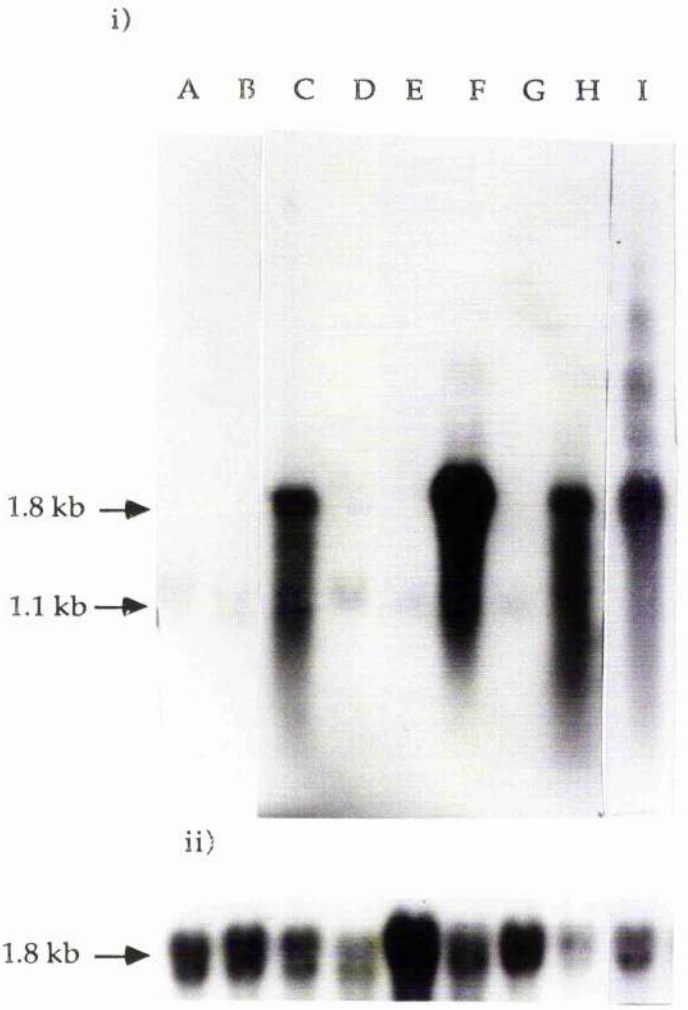


Figure 3.11 Northern blot of total *A.nidulans* RNA isolated from strains grown in MM with 10 mM glutamate. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *crnA* cDNA fragment of pSTA1500 and ii) *A.nidulans actA* fragments. Lane A; wild type, lane B; *niaD* 21, lane C; *niaD* 26, lane D; *niaD* 42, lane E; *niiA* 17, lane F; *cnxE* 17, lane G; *crnA* 1, lane H; Δ509, lane I; *niaD* 5.

hypothesis that NR may play a role in the regulation of the nitrate cluster genes (1.4.7).

One or more strains having a mutation within the *niaD* or *cnxE* genes were tested for growth on MM with ammonium or nitrate and MM with 200 mM chlorate and either proline, urea or glutamate as the nitrogen source (2.1.3.a). All strains were capable of growth on ammonium but only the wild type supported growth on nitrate, suggesting that all *niaD* and *cnxE* gene mutations result in a strain with a non-functional NR enzyme. The result verifies that several mutations at the *niaD* locus and one at the *cnxE* locus allow growth on chlorate (Table 3.6).

A sample of mutant strains was selected to illustrate the effect the NR enzyme has on the regulation of *crnA* gene expression. All strains were normal for the induction and repression of the *crnA* gene, with the exception of *crnA*1 which resulted in the loss of *crnA* mRNA, indicating that the mutation is possibly within the upstream regulatory sequences. Only the chlorate resistant strains *niaD*26, *cnxE*17, Δ509 and *niaD*5 allowed the constitutive synthesis of the *crnA* transcript (Fig 3.12 i lanes C, F, H and I respectively). In contrast, all other strains, excluding *crnA*1, exhibited normal *crnA* gene expression. No significant difference between *actA* transcript levels in each lane is observed (Fig 3.11 ii) demonstrating essentially equal RNA transfer.

The *niaD*5 strain has a mutation within the upstream regulatory sequences of the *niaD* gene and the complete intergenic region is lost in the *niiA-niaD* Δ509 strain (Fig 1.6). The results in section 3.4.4. illustrate that these particular mutant strains lose the ability to synthesize a *niaD* message and extrapolating, the NR enzyme. This implies that NR is not a pre-requisite for induction of the *crnA* gene. That certain *niaD* and *cnxE* mutations allow the constitutive expression of the *crnA* gene supports

the notion that the auto-regulatory role played by a functional NR is one of repression (1.4.7) (Cove and Pateman, 1969).

3. 4. 5. Temporal aspects of *crnA* gene expression

Previous results (1.4.2) implied that the *crnA* gene may be developmentally regulated, active only in conidia and young mycelia (Brownlee and Arst, 1983). It was suggested that nitrate may be transported by a second permease in *A.nidulans* which is only expressed in older mycelia. Initially the smaller, 1.1 kb message detected by the *crnA* cDNA on Northern blots, was a possible candidate for this (3.4.0).

To investigate this developmental phenomenon in terms of the level of mRNA accumulation, RNA was extracted from wild type mycelia grown under similar conditions to those employed by Brownlee and Arst (1983). Incubation was at 37° C in MM with urea as the nitrogen source. Expression of the *crnA* gene was induced by adding nitrate to a concentration of 10 mM, 100 min before harvesting. The results (Fig 3.12 i) show that *crnA* mRNA is abundant in 9 and 16 h mycelia (lanes A and B, respectively), the level decreasing in 20 (lane C) and 24 h (lane D) mycelia. That an *actA* transcript (Fig 3.12 iii) is detected at 20 and 24 h indicates that the decrease in *crnA* message level is real and not a result of unequal transfer. This is in contrast to previous results which demonstrated a decrease in nitrate transport at 16 h.

Glucose has been shown to be necessary for the expression of the *crnA* gene (3.4.2). It was considered that after 20 h at 37° C the MM may have a lower than optimal glucose concentration. Therefore, the above experiment was repeated adding glucose to 1% in addition to the nitrate, 100 min before harvesting. The results show (Fig 3.12 ii) that even in 30 h old mycelia (lane C) the *crnA* gene can be induced (RNA transfer being

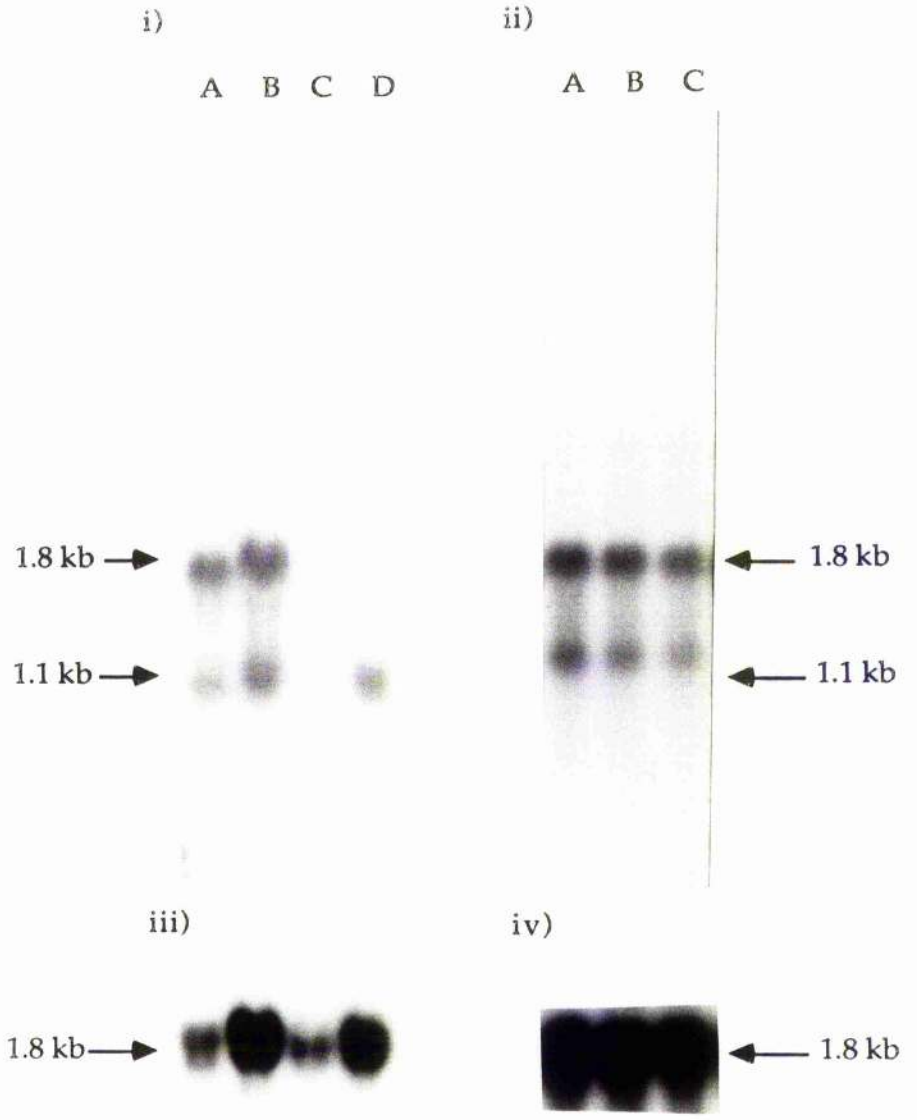


Figure 3.12 Northern blots of *A.nidulans* wild type total RNA isolated from mycelia grown initially in MM with 5 mM urea. Subsequently, i) and iii) nitrate was added to 10 mM and incubated for a further 100 minutes after the initial time periods of, lane A; 10 hours, lane B; 15 hours, lane C; 20 hours and lane D; 30 hours. ii) and iv) nitrate was added to 10 mM and glucose to 1% and incubated for a further 100 minutes after the initial time periods of, lane A; 10 hours, lane B; 20 hours and lane C; 30 hours.

20 µg of RNA was electrophoresed on 1.2% formaldehyde denaturing agarose gels, transferred to nylon membrane and hybridised to the hexaprime labelled, i) and ii) *crnA* cDNA fragment from pSTA1500 or, iii) and iv) *A.nidulans actA* fragments.

equal, Fig 3.12 iv), demonstrating that the loss of a *crnA* message in the previous experiment was most likely due to a limiting level of glucose. As a consequence, it could be assumed that the *crnA* gene is not developmentally regulated. However, this assumption is not supported by the decrease in the level of actin message in glucose limiting conditions (Fig 3.12 iii lanes C and D). Previous results demonstrated a significant decrease in the *actA* transcript when glucose was absent from the media (Fig 3.10 ii). Perhaps the difference in *actA* gene expression is reflected by the difference in glucose concentration. When glucose is limiting it is possible that only essential genes are transcribed, i.e. *actA* (Fig 3.12 iii lanes C and D), compared with *crnA* (Fig 3.12 i lanes Cand D), whereas when glucose is absent *actA* gene expression decreases significantly (Fig 3.10).

3. 5. 0. REGULATION OF *niiA* and *niaD* GENE EXPRESSION

The regulation of *niiA* and *niaD* gene expression is known to be by induction in the presence of nitrate and a functional *nirA* gene product, with repression mediated by the *areA* gene product in the presence of ammonium or glutamine (1.4.7). However, the level of *niiA* and *niaD* gene expression had not been determined.

To ensure *niiA* and *niaD* mRNA detection 10 µg of total RNA extracted from mycelia grown under varying conditions was electrophoresed on a 1.2% formaldehyde denaturing agarose gel as before (3.4.0). After Northern blotting the filters were hybridised against either the hexaprime radio-labelled (2.5.5)*Xba*I fragment of pSTA1 (Fig 2.2), from within the coding region of the *niaD* gene, or the complete *EcoR*I fragment of pSTA3 (Fig 2.2) containing part of the *niiA* gene.

That the *niaD* and *niiA* mRNAs could be detected with only 10 μg of total RNA suggests higher abundance than the *crnA* mRNA. Additionally, due to stronger hybridisation it can be inferred that the *niiA* mRNA is more abundant than the *niaD* mRNA.

3. 5. 1. The influence of different nitrogen sources on *niiA* and *niaD* gene expression

The results (Figs 3.13 i and 3.14 i) clearly show that *niiA* and *niaD* gene expression is regulated as described (1.4.7) and occurs at the level of mRNA accumulation. Nitrate and nitrite (lanes A and B, respectively) both induce the expression of the *niiA* and *niaD* genes. Incomplete transfer of RNA from nitrate grown mycelia, as demonstrated by the actin blot (Fig 3.13 ii lane A), suggests that the expression of the *niaD* gene is greater, relatively than shown in Fig 3.14 i, lane A. In contrast, ammonium (lane C) in addition to glutamate (lane D) and proline (lane E) do not induce the expression of the *niiA* and *niaD* genes. That ammonium is responsible for repression is confirmed by a comparison of lanes F and G. When nitrate and proline are present together (lane F) both genes are expressed. In contrast, the introduction of ammonium, results in the loss of the *niiA* and *niaD* transcripts, overriding nitrate-mediated induction (lane G). Generally, (except as already discussed) RNA loading and transfer was essentially similar in each lane, as indicated by the *actA* transcript levels (Fig 3.13 iii and 3.14 iii).

3. 5. 2. Regulation of *niiA* and *niaD* gene expression by the *nirA* and *areA* genes

The results (Figs 3.13 ii and 3.14 ii) demonstrate that the *nirA* gene is involved in the induction of *niiA* and *niaD* gene expression. The pleiotropic loss of function mutant, *nirA*1, fails to induce the expression

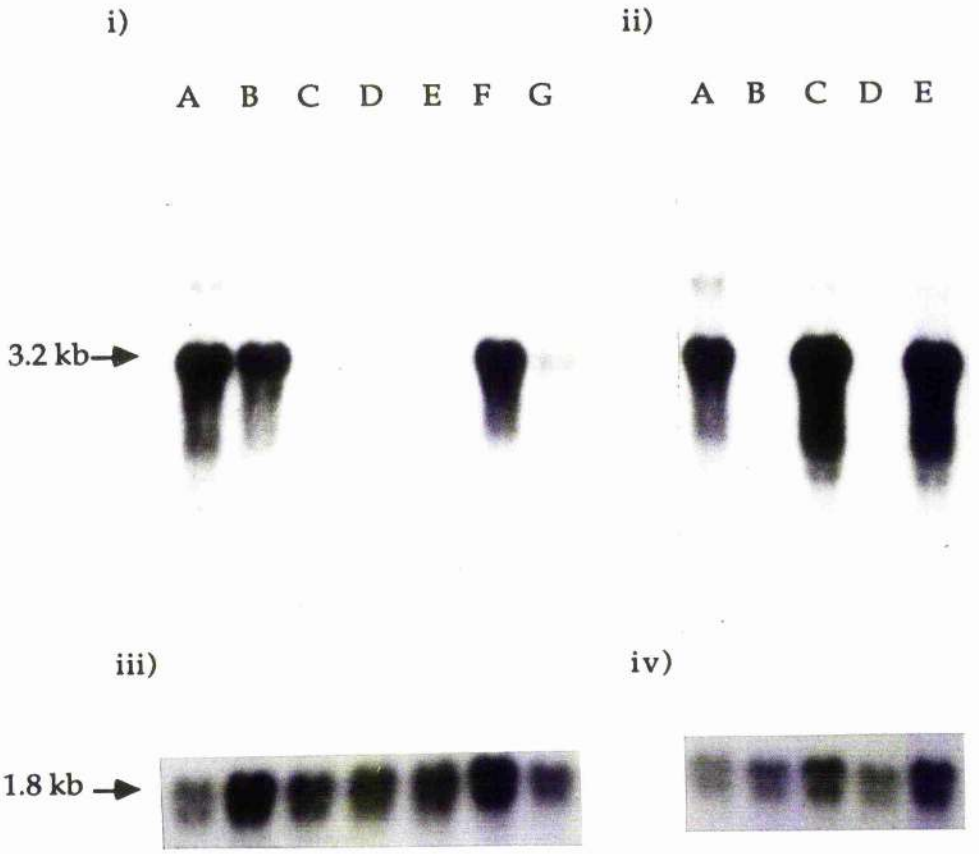


Figure 3.13 Northern blot of *A.nidulans* total RNA. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) and ii) *EcoR* I fragment from pSTA3, or iii) and iv) *A nidulans actA* fragments.

i) and iii) wild type grown in MM with, lane A; 10 mM nitrate, lane B; 10 mM nitrite, lane C; 10 mM ammonium, lane D, 10 mM glutamate, lane E; 10 mM proline, lane F; 10 mM nitrate and 10 mM proline, lane G; 10 mM nitrate and 10 mM ammonium.

ii) and iv) lane A; wild type grown in MM with 10 mM ammonium and transferred to 10 mM nitrate, lane B; *nirA* 1 grown under similar conditions to the wild type, lane C; *nirA* ^{c} grown in 10 mM glutamate, lane D; *areA* 19 grown under similar conditions to the wild type, lane E; *xprD* 1 grown in MM with 10 mM nitrate and 10 mM ammonium.

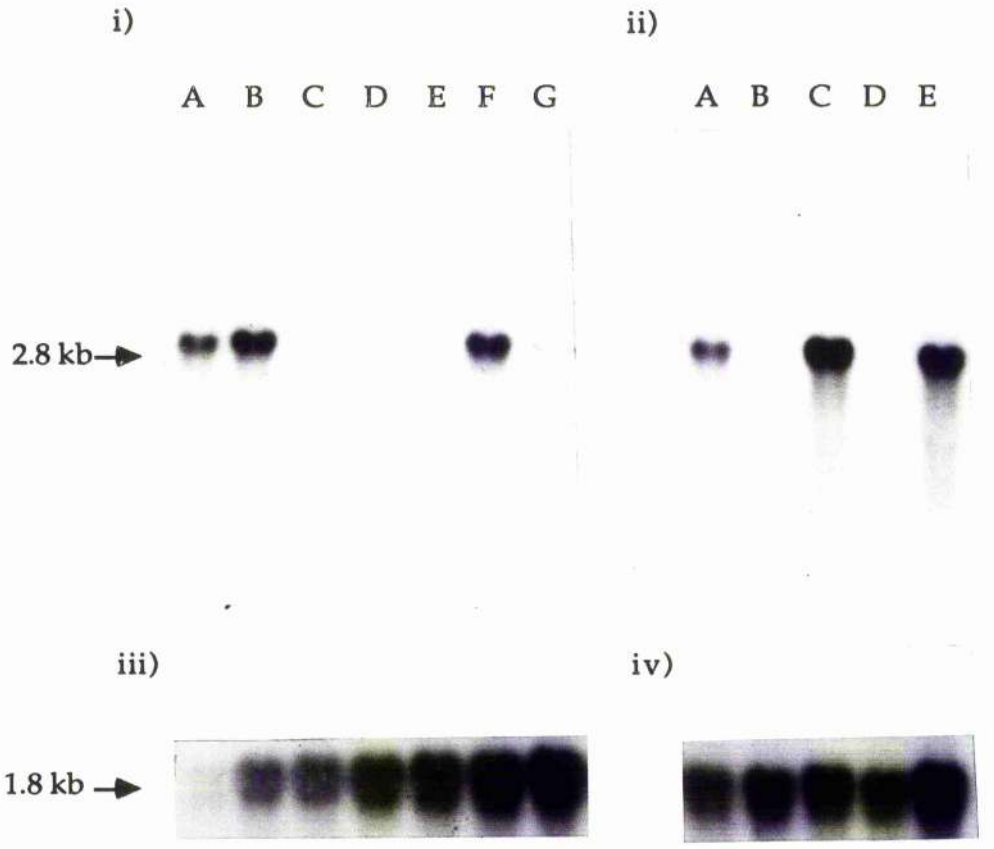


Figure 3.14 Northern blot of *A.nidulans* total RNA. 20 μg was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) and ii) *Xba* I fragment of pSTA1 or iii) and iv) *A.nidulans actA* fragments.

i) and iii) wild type grown in MM with, lane A; 10 mM nitrate, lane B; 10 mM nitrite, lane C; 10 mM ammonium, lane D; 10 mM glutamate, lane E; 10 mM proline, lane F; 10 mM nitrate with 10 mM proline, lane H; 10 mM nitrate with 10 mM ammonium.

ii) and iv) lane A; wild type growm in MM with 10 mM ammonium and transferred to 10 mM nitrate, lane B; *nirA* 1 grown under similar conditions to the wild type, lane C; *nirA* ^{c} grown in MM with 10 mM glutamate, lane D; *areA* 19 grown under similar conditions to the wild type and lane E; *xprD* 1 grown in MM with 10 mM nitrate and 10 mM ammonium.

of the *niiA* or *niaD* genes despite the presence of the inducer, nitrate (lane B) (mycelia were grown in 10 mM ammonium and transferred to nitrate), suggesting a requirement for the *nirA* gene in *niiA* and *niaD* gene expression. To confirm this, the *niaD* and *niiA* genes are shown to be constitutively expressed in the *nirA*^{c}1 mutant background (lane C), i.e. both genes are expressed when glutamate is the sole nitrogen source, contrasting sharply to the results observed with the wild type strain (Fig 3.13 i and 3.14 i).

The *areA* loss of function mutant strain, *areA*19, is, in parallel to the *nirA*1 strain, unable to procure the expression of the *niiA* and *niaD* genes under inducing conditions (Fig 3.13 and 3.14 ii, lane D). This suggests that the *areA* gene is essential for *niiA* and *niaD* gene expression, supported by the observation that derepression of the *niiA* and *niaD* genes occurs in the *xprD*1 mutant background (Fig 3.13 and 3.14 ii lane E), i.e. both genes are expressed when ammonium and nitrate are available together as the nitrogen source. Nitrate induction and ammonium repression (when nitrate and ammonium are present as sole nitrogen sources respectively) of *niiA* and *niaD* gene expression in the *nirA*^{c} and *xprD* mutant backgrounds is similar to that of the wild type (Fig 3.13 and 3.14 i).

Again, these differences in *niiA* and *niaD* transcript levels are shown to be real by the similarity of *actA* transcript levels in each lane (Fig 3.13 and 3.14 iv), demonstrating equal transfer of RNA.

3. 5. 3. Regulation of *niiA* and *niaD* gene expression by NR

As discussed previously (3.4.4), the NR enzyme is proposed to exert a regulatory effect on the *niiA* and *niaD* genes. The mutant strains *cnxE*17, *niaD*5, *niaD*26 and *niiA-niaD* Δ509 exhibiting chlorate resistance were shown to be capable of constitutive *crnA* gene expression.

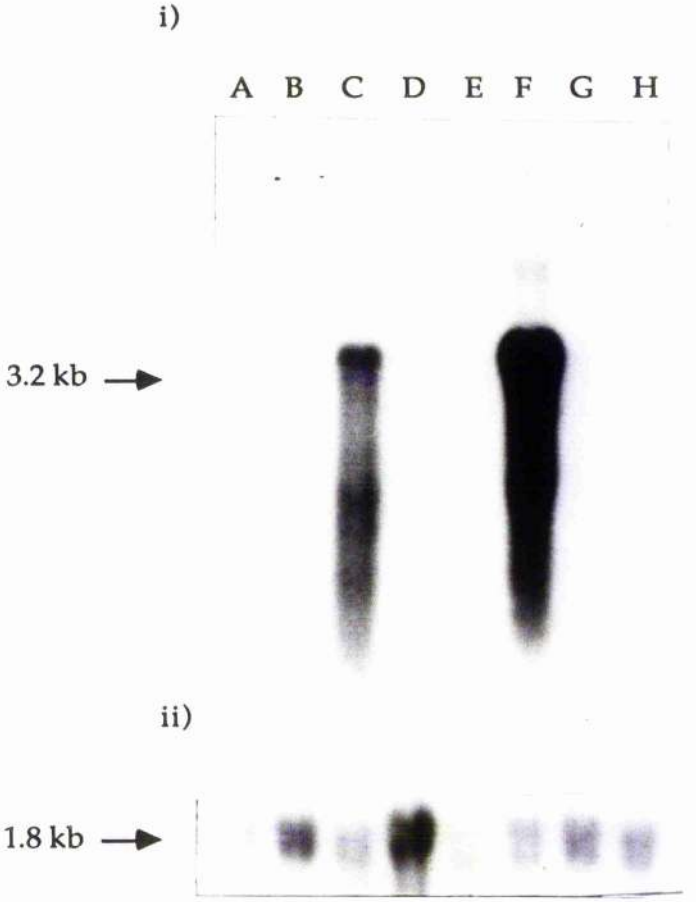


Figure 3.15 Northern blot of *A.nidulans* total RNA isolated from strains grown on 10 mM glutamate. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *EcoR* I fragment from pSTA3 or ii) *A.nidulans actA* fragments. Lane A; wild type, lane B; *niaD* 21, lane C; *niaD* 26, lane D; *niaD* 42, lane E; *niiA* 17, lane F; *cnxE* 17, lane G; *crnA* 1 and lane H; Δ509.

The expression of the *niiA* gene was demonstrated to be regulated in a similar way to the *crnA* gene in many of the mutant backgrounds, although *niiA* gene expression in the *niaD*5 strain was not examined. The difference between *crnA* and *niiA* gene expression in the various strains is observed with the deletion mutant *niiA-niaD* Δ509 (Fig 3.15). This mutation lacks the upstream regulatory sequences for the *niiA* gene (Fig 1.6) and as expected, does not allow *niiA* gene expression despite constitutive *crnA* gene expression.

Table 3.6 *A.nidulans* wild type and various mutant strains grown on MM with nitrate in addition to MM with 200 mM chlorate and either glutamate, urea or proline as the nitrogen source. + indicates good growth, +/- indicates a small amount of growth and - indicates no growth. Media is described in section 2.1.3.a and the strains are listed in section 2.1.4.a.

Strain	Media			
	MM + Nitrate	MM+200 mM Chlorate and :		
		Glutamate	Urea	Proline
W.T. (*biA* 1)	+	-	-	-
G834 (*nirA* 1)	-	+	+	+
B400 (*niaD* 42)	-	-	+/-	+/-
B556 (*niaD* 118)	-	-	+	+/-
B366 (*niiA* 17)	-	-	-	-
B125 (*niaD* 26)	-	+	+	+
B352 (*niaD* 21)	-	-	+	+/-
G059 (*cnxE* 17)	-	+	+	+
B286 (*niaD* 5)	-	+	+/-	+/ -
B344 (*niaD* 10)	-	+	+	+
B348 (*niaD* 16)	-	+	+	+
B351 (*niaD* 20)	-	-	+/-	+/-
B397 (*niaD* 40)	-	-	+/-	+/-
B468 Δ509	-	+	+	+

To extend this work and with the eventual aim of identifying a region of the *niaD* gene product responsible for the regulatory function it was necessary to examine further *niaD* mutant strains for chlorate resistance (Table 3.6) and subsequently, constitutive expression of the *niaD* gene.

Many *niaD* mutant strains are available but although the extent of various deletions have been mapped (Unkles pers. comm.), in few cases is the exact location of a point mutation known. An attempt to locate the position of some of these mutations by restriction analysis was conducted to enable a more accurate diagnosis of the region

---

Table 3.7 The sizes of the restriction fragment lengths resulting from a *Taq*I digest of the *niaD* gene. An asterisk indicates that the fragment results from a *Taq*I site within an intron. Numbers in bold type are those that may be identified after electrophoresis through a 2% agarose gel. Brackets indicate fragments that migrate together. A total of 20*Taq* I restriction sites within exon sequences and 1 within intron sequences releases 22 fragments from the *niaD* gene:

1	**765 bp**	93
2	**>450**	75 *
3	**(301**	64
	(299 *	48
4	**(221**	32
	(210	30
5	**(169**	27
	(165	27
6	**117**	26
	108	24
	106	15

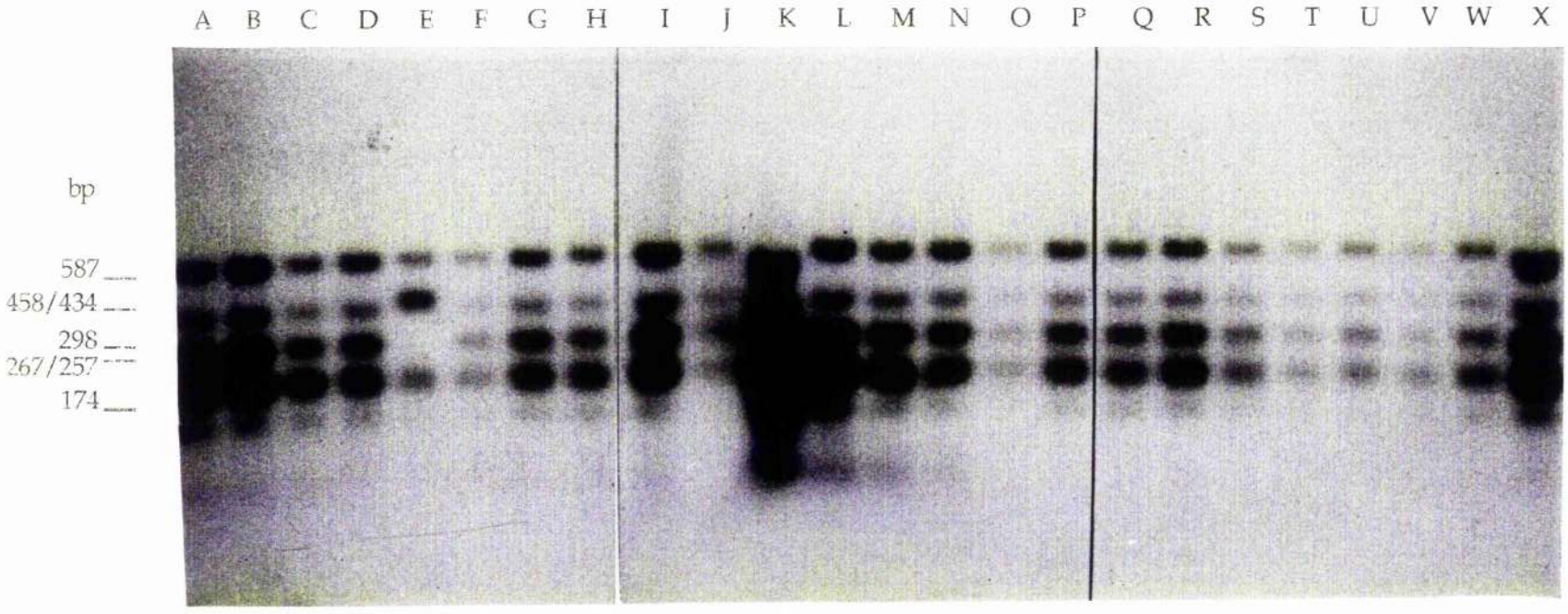


Figure 3.16 Southern blot of *A.nidulans* genomic DNAs digested to completion with endonuclease *Taq* I. Approximately 10 µg of DNA was electrophoresed on a 2% agarose gel, transferred to nylon membrane and hybridised to hexaprime radio-labelled pILJ141. DNA was isolated from the strains: lane A; *niaD* 1, lane B; *niaD* 8, lane C; *niaD*28, lane D; *niaD*29, lane E; *niaD*118, lane F; *niaD* 119, lane G; *niaD*120, lane H; *niaD*124, lane I; *niaD*127, lane J; *niaD*17, lane K; *niaD*25, lane L; *niaD*123, lane M; *niaD*31, lane N; *niaD* 35, lane O; *niaD* 54, lane P; *niaD*170, lane Q; *niaD* 34, lane R; *niaD*24, laneS; *niaD*2, lane T; *niaD*115, lane U; *niaD*15, lane V; *niaD*122, lane W; *niaD*171, lane X; wild type.

of the *niaD* gene, or its product, involved with autoregulation. The restriction enzyme chosen for use was *Taq*I. The recognition site of *Taq*I consists of only four bases and therefore occurs relatively frequently throughout DNA generally and the *niaD* cDNA. Only *Taq*I sites within the cDNA were considered since intron sequences do not contribute to the polypeptide and hence base changes within these regions will not affect the structure of the NR enzyme. A total of 20 *Taq*I restriction sites allows the investigation of 80 nt, 3% of the *niaD* cDNA. Unfortunately not all the *niaD* restriction fragments obtained from a complete *Taq*I digest can be resolved on a 2% agarose gel (Table 3.7), reducing the number of bases examined to 20.

The restricted DNA from 23 mutants and the wild type was electrophoresed through a 2% agarose gel alongside *Hae*III cut pUC DNA molecular weight markers (Fig 2.5). DNA was transferred to nylon membrane by Southern blotting (2.5.4.a) and hybridised at 64° C (2.5.6.b) to hexaprime radiolabelled *Hind*III cut pILJ141 (Fig 2.2), housing the whole of the *niaD* gene.

The results (Fig 3.16) illustrate a change in the hybridisation profile of the strain *niaD*118 (lane E). The third band of 301 bp is absent, and the second band shows an increase in size, inferring a disruption of the *Taq*I site at 1949 nt in the *niaD* gene. The hybridisation profile is misleading as the band of 299 bp is also missing in the *niaD*118 strain. It is difficult to explain this although the mutation is not a nucleotide change at either *Taq*I site creating the 299 bp fragment. One of these *Taq*I sites falls within an intron and disruption of the other would result in a fragment of 1 kb. A fragment of this size was not detected by Southern blotting.

None of the remaining mutant strains could be characterised by this method, presumably because no other mutations fall within a *Taq*I recognition site. The location of point mutations is therefore restricted by

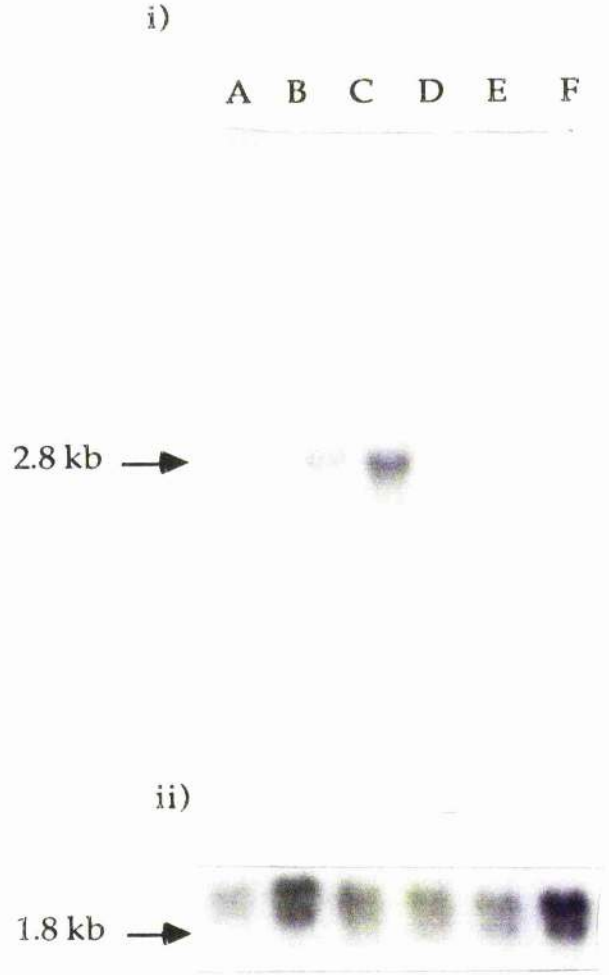


Figure 3.17 Northern blot of *A.nidulans* RNA isolated from various strains grown on MM with 10 mM glutamate. 20 µg of total RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *Xba* I fragment of pSTA1 and ii) *A.nidulans actA* fragments. Lane A; *niaD* 5, lane B; *niaD* 10, lane C; *niaD* 16, lane D; *niaD* 20, lane E; *niaD* 40, lane F; *niaD* 118

this method. A more useful approach may be to generate a new set of point mutations by site-directed mutagenesis. Consequently, in addition to the *niaD*118 mutant strain and five others chosen arbitrarily the regulation of *niaD* gene expression by NR was examined in those mutant strains used previously to study the regulation of *crnA* and *niiA* gene expression.

The mutant strains *niaD*10, *niaD*16 (Fig 3.17 i lanes B and C), *niaD*26 and *cnxE*17 (Fig 3.18 i lanes C and F) demonstrated the ability to constitutively express the *niaD* gene. These strains are all chlorate resistant (Table 3.6). The remaining two chlorate resistant strains, *niaD*5 and *niiA-niaD* Δ509 contain mutations within the upstream regulatory sequences of the *niaD* gene (Fig 1.6), preventing its expression. The strains *niaD*21 and *niaD*42 (Fig 3.18 i lanes A and D) did not constitutively express the *niaD* gene and are found to be chlorate sensitive (Table 3.6).

That the *crnA*1 and *niiA*17 strains were not capable of constitutive *niaD* gene expression suggests that only the *niaD* gene product is involved with the autoregulation of the nitrate gene cluster. It is important to note that nitrate induction and nitrogen metabolite repression of *niaD* gene expression, as judged by transcript analysis, was normal, i.e. as for the wild type (Fig 3.14 i), in all mutant strains except *niaD*5 and Δ509. These two strains may therefore have disruptions within the 5' regulatory sequences.

An *actA* transcript is observed in each lane (Fig 3.17 and 3.18 ii), after the Northern blot was stripped of radioactivity and re-probed with a fragment of the *A.nidulans actA* gene. This indicates that the observed differences in *niaD* transcript levels (Fig 3.17 and 3.18 i) are real and not a result of unequal transfer.

Consequently, some mutations within the *niaD* and *cnxE* genes, allow the constitutive expression of the *niiA*, *crnA* and *niaD* genes. The

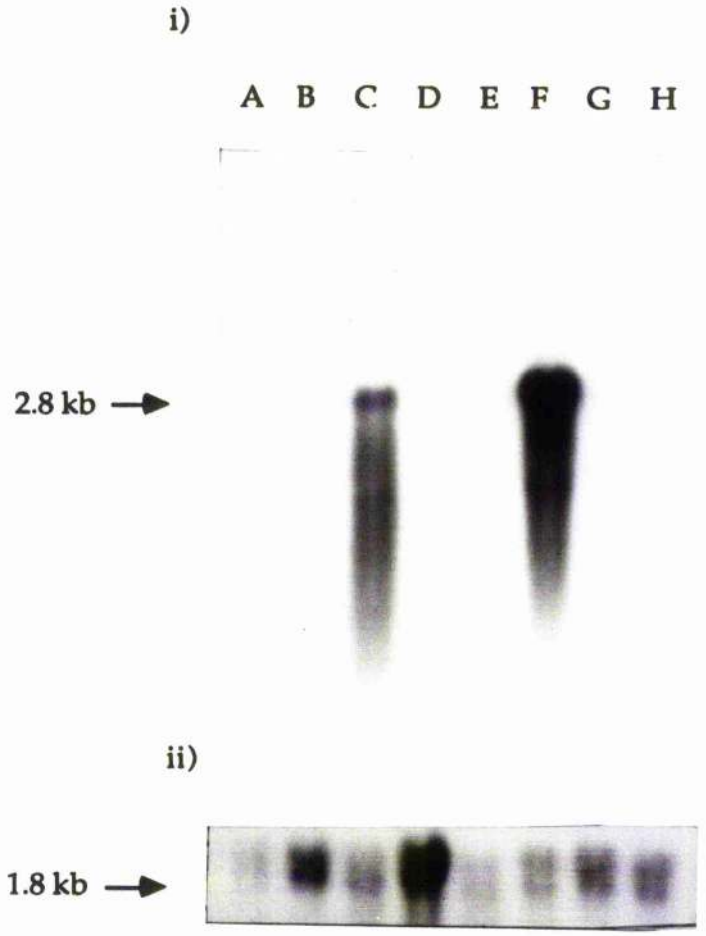


Figure 3.18 Northern blot of *A.nidulans* total RNA isolated from strains grown on 10 mM glutamate. 20 µg of RNA was electrophoresed on a 1.2% formaldehyde denaturing agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled, i) *Xba* I fragment from pSTA1 or ii) *A.nidulans actA* fragments. Lane A; wild type, lane B; *niaD* 21, lane C; *niaD* 26, lane D; *niaD* 42, lane E; *niiA* 17, lane F; *cnxE* 17, lane G; *crnA* 1 and lane H; Δ509.

NR enzyme would therefore appear to be necessary for the repression of the nitrate gene cluster when nitrate is not available. The *nirA* gene product is required for the expression of the *crnA*, *niiA* and *niaD* genes in the presence of nitrate (3.4.3 and 3.5.2). It is possible that the *nirA* gene is expressed constitutively, dimerising with the NR enzyme in a nitrate-free environment to prevent expression of the nitrate gene cluster. Alternatively, the NR enzyme may prevent expression of the *nirA* gene by interaction with upstream *cis*-acting regulatory sequences. Either possibility requires the NR enzyme to have a structural motif enabling interaction with either another protein or DNA sequences. These motifs may be expected to occur within areas of the *niaD* gene in which mutations lead to chlorate resistance. However, the data from this study implies that chlorate resistance and constitutive gene expression result from small deletion mutations. Point mutations within the area encompassed by these deletions do not appear to alter normal gene activity. This suggests that extensive changes within the tertiary structure of the NR enzyme are necessary before any change in *crnA*, *niiA* and *niaD* gene expression is seen. Alternatively, it is possible that a motif necessary for protein (Fig 1.2) or DNA-binding occurs within a region of the *niaD*26 mutation outwith the *niaD*21 and *niaD*42 mutations, for example (Fig 1.6). Sequence comparisons with examples of leucine zippers, homeodomains, helix-loop-helices and zinc fingers (1.3.1) did not positively identify a similar motif within the *niaD* gene product. In addition, it was difficult to find any region of the *niaD* encoded polypeptide with a higher than average concentration of either glutamine, proline or acidic residues (1.3.1).

3. 6. 0. UPSTREAM SEQUENCE ANALYSES OF *niiA* and *niaD* GENES

It has been demonstrated that the *crnA*, *niiA* and *niaD* genes of *A.nidulans* are regulated in a similar manner, requiring both functional *nirA* and *areA* genes for their expression (3.4.3 and 3.5.2). It is plausible then that the genes of the nitrate cluster may share common upstream regulatory sequences with a *cis*-acting regulatory function, interacting with the *nirA* and *areA* gene products (1.4.7). The availability of the *A.oryzae* and *A.niger niiA-niaD* intergenic region sequences allowed a direct comparison with the motifs common to the 5' non-coding sequences of the *A.nidulans niiA* and *niaD* genes. The regulation of the *niiA* and *niaD* genes in these three species of *Aspergillus* is assumed to be similar especially in the light of recent evidence suggesting similar regulation for the nitrate assimilatory genes of *A.nidulans* and its more distant relative *N.crassa* (Davis and Hynes, 1987, section 3.6.0 and Fu and Marzluf, 1990).

3. 6. 1. Comparison of the 5' non-coding sequences of the *A.nidulans niiA* and *niaD* genes.

The 1262 bp *niiA-niaD* intergenic region (Johnstone *et al*, 1990) was divided in the centre resulting in approximately 600 bp upstream of the ATG of each gene. A computer-aided comparison of the sense strand of the upstream sequence of each gene identified eight short sequences occupying similar relative positions (Table 3.8) within 400 bp of the translational start site. The position of the motifs should ideally be related to the transcriptional start site since their proposed effect is on transcription and not translation. However, the location of transcriptional initiation had not been established for the *niaD* gene.

Table 3.8 Motifs common to the 5' non-coding regions of the *A.nidulans niiA* and *niaD* genes. The distance indicating the position from the translational start site of each motif excludes the ATG and the first nucleotide in each motif.

Motif	Position from translational start site	Sequence	Identity	Length(nt)
1	*niiA* -23	TCAGACCCCCT	64%	11
	niaD -29	TCTGTACCTCT		
2	*niiA* - 98	TGACTCGTTGTGCCACACTAT	57%	21
	niaD -104	AGAACCCGTGCCCTATACTAT		
3	*niiA* -126	TCACCAAGG	67%	9
	niaD -137	TCCCCCACG		
4	*niiA* -169	GTCATTGGCCCATTT	67%	15
	niaD -187	GTTTTCGTCTCATTT		
5	*niiA* -198	TCGCTGATTCG	82%	11
	niaD -216	TCT-TGATTTG		
6	*niiA* -259	TC-CAGATCCTAAA	64%	14
	niaD -275	TCTCCGAGGCTGAA		
7	*niiA* -305	CCAACATCATGCTTATCGC	58%	19
	niaD -321	CCAGCATATTGCCTGCTGT		
8	*niiA* -352	AGTGCATCGAAGCCATCCACGAT	74%	23
	niaD -333	AGTGGATCGTAGG-ATCCAGCAT		

The motifs range from 9 to 23 nt in length and share between 57% and 82% of nucleotides in common. It is questionable whether these motifs serve any real purpose, especially as a recognition site for a *trans*-acting factor, as their sequences are not homologous between the two genes. However, the motifs common to the 5' non-coding regions of the *qut* (Hawkins *et al*, 1988) and the *qa* (Baum *et al*, 1987) genes also vary in

sequence, although they do have a conserved trend. For example, the *qa* consensus has partial dyad symmetry and is therefore capable of forming a hairpin loop (1.2.1).

The significance of all eight motifs is again questionable since it is believed that only the *nirA* and *areA* gene products bind to *cis*-acting sequences of the *niiA* and *niaD* genes. Perhaps more than one of these motifs constitute a recognition element for a single *trans*-acting factor. Alternatively, the presence of several of the sequences may be entirely fortuitous.

A computer-aided search for the presence of the eight motifs within the sequence 5' to the *crnA* translational start site revealed regions similar to only motifs 4 and 5 (3.3.2). That motif 5 of the *crnA* gene shares 73% sequence identity with that of the *niiA* and *niaD* genes perhaps increases the possibility that its role is functional.

3. 6. 2. Sequence comparison of the *niiA-niaD* intergenic region of *A.nidulans, A.oryzae* and *A.niger*

The nucleotide sequence of the *niiA* and *niaD* genes of *A.oryzae* and *A.niger* was determined by Dr. Shiela Unkles. The contiguous arrangement of the two genes and the fact that they are divergently transcribed shows similarities with *A.nidulans*. In contrast, the size of the *A.oryzae* and *A.niger* intergenic regions are 1684 bp and 1653 bp respectively, approximately 400 bp larger than the equivalent region in the *A.nidulans* genome. This difference may suggest a closer phylogenetic relationship between *A.oryzae* and *A.niger*, the *A.nidulans* genome having lost the unnecessary nucleotides during the course of evolution.

Nitrate assimilation in the three species may be assumed to be regulated in a similar manner due to their close phylogenetic relationship. In support of this the *N.crassa* regulatory genes *nit-2* and

nit-4 have been shown to complement *A.nidulans areA* and *nirA* loss of function mutations respectively, suggesting that the regulatory gene products of *N.crassa* and *A.nidulans* recognise similar *cis*-acting sequences.

The *niiA-niaD* intergenic regions of *A.oryzae* and *A.niger* were divided in a similar manner to that of *A.nidulans* (3.6.1) and a computer-aided search for the eight motifs common to the *niiA* and *niaD* 5' non-coding sequences was conducted. Several regions of 70% similarity with the motifs of 15 nt or less, i.e. 1, 3 and 5 can be identified although none have greater identity. In particular, motif 3 is repeated many fold, rendering an interpretation meaningless. (For this reason motif 3 has not been included in Table 3.9). Areas of 70% identity with the longer motifs, 2, 7 and 8, could not be identified and on decreasing the percentage similarity several regions were again revealed. Sequences 50% similar to motif 2 for the *A.niger niaD* gene and motif 7 for the *A.oryzae niaD* gene could not be determined. The motifs selected for illustration in Table 3.9 are within 500 bp of the ATG site of each gene and occupy a similar relative position to those of the *A.nidulans niiA* and *niaD* genes, where possible.

The motifs from the six genes (*niiA* and *niaD* of *A.niger*, *A.oryzae* and *A.nidulans*) were compared and a consensus sequence deduced for each (Table 3.9). The percent identity of each motif to the consensus is indicated (Table 3.9). In many cases the percent identity may be meaningless as, particularly for motifs 1, 2, 4, 7 and 8 the consensus is very general with a high concentration of Y, R and N, any pyrimidine, any purine and any nucleotide respectively..

The *niiA-niaD* intergenic regions of *A.oryzae*, *A.niger* and *A.nidulans* and the 5' flanking sequences of the *crnA* gene were also

Table 3.9 Motifs common to the 5' non-coding sequences of the *niiA* and *niaD* genes of *A.nidulans, A.oryzae* and *A.niger*. A consensus sequence deduced for each motif is shown. The number representing the distance of each motif from the translational start site excludes the ATG and the first nucleotide of the motif.

Motif	Species	Gene	Position from ATG	Sequence	Similarity with Consensus
1	*A.nidulans*	*niaD*	-29	TCTGTACCTCT	73%
		niiA	-23	TCAGACCCCCT	100%
	A.oryzae	*niaD*	-35	TCCCCTCCCTC	64%
		niiA	-75	ACATACTCCTT	64%
	A.niger	*niaD*	-28	TTGGCACCTCG	64%
		niiA	-87	TCGGGTTCCCT	91%
	Consensus			TCRGNYCCCCT	
2	*A.nidulans*	*niaD*	-104	AGAACCCGTGCCCTATACTAT	95%
		niiA	-98	TGACTCGTTGTGCCACACTAT	90%
	A.oryzae	*niaD*	-152	GTAACACGTACTGCGTGCCAT	67%
		niiA	-181	TGATTCGATAGTCAAACATCT	57%
	A.niger	*niiA*	-211	AGTCCCACGGAGAAACACAAT	67%
	Consensus			NGAYCCNNTGYYCYAYACTAT	
4	*A.nidulans*	*niaD*	-187	GTTTTCGTCTCATTT	100%
		niiA	-169	GTCATTGGCCCATTT	80%
	A.oryzae	*niaD*	-18	ATTTTACTTTCATTG	60%
		niiA	-339	ATCCCTGACCCAATC	60%
	A.niger	*niaD*	-109	TTGCTCCTCTCATCT	73%
		niiA	-238	GTCATAGTCACCCCT	60%
	Consensus			GTCYTYGTCTCATTT	

5	*A.nidulans*	*niaD*	-216	TCT-TGATTTG	100%
		niiA	-198	TCGCTGATTCG	73%
	A.oryzae	*niaD*	-103	TCC-TTATTTG	91%
		niiA	-195	AGG-TGATTCG	64%
	A.niger	*niaD*	-174	TCT-GGTTTTT	73%
		niiA	-301	CCC-TGATTAC	73%
	Consensus			TCY-TGATTTG	
6	*A.nidulans*	*niaD*	-275	TCTCCGAGGCTGAA	93%
		niiA	-259	TC-CAGATCCTAAA	71%
	A.oryzae	*niaD*	-230	TCTCCGAGGCGTTA	93%
		niiA	-210	TCGATGACCATCAA	50%
	A.niger	*niaD*	-262	TCTCCGTGGACTTA	79%
		niiA	-372	TCCCAGTTCTTATC	50%
	Consensus			TCTCCGAGGCTNTA	
7	*A.nidulans*	*niaD*	-321	CCAGCATATTGCCTGCTGT	89%
		niiA	-305	CCAACATCATGCTTATCGC	95%
	A.oryzae	*niiA*	-158	ACGACCTCATGGTGTCCAC	63%
	A.niger	*niaD*	-44	CCTG-ATATTGTCTCATGC	79%
		niiA	-323	CCTCTATGATCTGCATGGC	63%
	Consensus			CCRRCATRATGYYTRYYGC	
8	*A.nidulans*	*niaD*	-333	AGTGGATCGTAGG-ATCCAGCAT	78%
		niiA	-352	AGTGCATCGAAGCCATCCACGAT	91%
	A.oryzae	*niaD*	-350	AAGCGGTGGACGCCGTCCAGCCT	65%
		niiA	-276	ATTGCCGATAACCCGTCCACATC	65%
	A.niger	*niaD*	-370	AACTCATTGTTCGAATCCACCGG	65%
		niiA	-497	GATTCATGCAGCCACTCCCCAAT	65%
	Consensus			AATGCATNGAANCCATCCACCAT	

compared with the 50 nt sequence common to the 5' non-coding sequences of the *A.nidulans* and *A.niger gpdA* genes (Punt *et al*, 1990), the *qa* consensus (Baum *et al*, 1987) and the *qut* consensus (Hawkins *et al*, 1988). A sequence with 70% similarity to the *qa* or *qut* consensus could be identified at least once within each of the *niiA-niaD* intergenic regions and upstream of the *crnA* gene. Little significance was drawn from this as sequence similarity between the identified motifs differed greatly and no partial or complete dyad symmetry was present in any of the sequences displaying identity with the *qa* consensus.

The *N.crassa nit*-2 gene product, for nitrogen metabolite repression, has been shown to recognise elements containing two copies of TATCTA, a similar sequence or its consensus, upstream of the *nit*-3 gene, encoding nitrate reductase. In addition, three such elements within the *niiA-niaD* intergenic region of *A.nidulans* have been demonstrated to be capable of *nit*-2 product binding, although their situation is unclear (Fu and Marzluf, 1990). The earlier complementation study (Davis and Hynes, 1987) and the above protein-binding analysis suggest that the *nit*-2 gene acts in a similar manner to the *areA* gene and that its product recognises *cis*-acting sequences of the *niaD* gene. It is likely that these TATCTA upstream elements are those that are recognised by the *areA* gene product.

A computer-aided search for sequences with 90% similarity to the motif TATCTA or its complement, within the *A.nidulans, A.oryzae* and *A.niger niiA-niaD* intergenic regions was conducted. Upstream of the *A.nidulans niaD* gene two TATCTA sequences occur within the region -77 to -97 nt and on the complementary strand between -487 and -526 nt (Fig 3.19). Two regions containing one copy of TATCTC in addition to one copy of a similar complementary sequence can be found upstream

A.nidulans niiA

```
-618                                                    A GTGATTTCAC CGGCGCTTTT

-597      TGGCAATTCC GATGACTCTC GATCGTGTCT CATAGATAGA GATAGGGCTT GATTATCGTT

-537      GATTCACTGT CCAATCAGAG CCTCGGAATC TCTTATCTCT GGTGGGAAGA AGGCGGTAGA

-477      TCATGGCACC ATTATGCCCA ATCAGAACGC TGCCCTGAGC CGTGGCCAGA CTTCCCAAAT

-417      ATCATCATTC GATCTCCAGC CCAACTCCGC GGAAATTCAG GCAGTGCATC GAAGCCATCC

-357      ACGATGATCC ACATCCACCA AACAATGACA GCGCCAACAT CATGCTTATC GCCGGCTCAT

-297      TATCGTCAAG CGGCTCTAGC CACACTCCAG ATCCTAAACC CCCGTTTCTC GTGGCGTTCA

-237      ATATCGATGA CGGTCGTCCA TTGCGAGATC GCTGATTCGC GTGTTACCAC ATAGTCATTG

-177      GCCCATTTAA TTATCTCATG CTCCGATGCC ACGGAGATAC AATCACCAAG GCGTTTGGTG

-117      ACTCGTTGTG CCACACTATG GATTCCCGCT ATATGAAGCT GGCAGCGTCG CCCGCGGCCG

 -57      GCTCTTGTGG GACTTGTGTT TTCTCAGACC CCCTGCTCAA TCACCGCGCC CGCCATCATG
```

Figure 3.19 5' non-coding sequences of the *A.nidulans*, *A.oryzae* and *A.niger niiA* and *niaD* genes. The translational start site is shown in bold. Numbers on the left-hand side represent the distance from the ATG site. Areas containing one or two copies of the consensus TATCTA or its complement are underlined.

A.nidulans niaD

-645 TTAAATAT TACATCGACT GTAATAGCCA TTGATCATAT TCTGGACAGT

-597 TCCCAGTTCC CGGAACATTC AGGATCGCGT AGGTGGCAGC AGTGACTGTA CACTATGAGC

-537 GGACTCCTAC CCAGATACGC TCACGCCATA ACCCTATTGC CACTAGATGA CAGTACGAGC

-477 CTCTGACCGT CTGGCTTGCC GTTGACCCGG GCTGCAGTAT TCTTGCTGTG TAATTACTTG

-417 ATATAGTACT GGTCAGGGCT ACACTTGCAG CCGGTTCGCG GCCCGCGGAG ATGGAGCCCT

-357 GAAGTGGATC GTAGGATCCA GCATATTGCC TGCTGTCCTT AGAATTAGAA AACGAAAATA

-297 CTCTGTTTCT CCGAGGCTGA AAGGCAGCGG GCGGCTGCCA ACTGTATTTC GTCTTATTTT

-237 GTCGTCTTTT GTCTTGATTT GGTCTTATTC CATCTGTTTT CGTCTCATTT CGTCAGATCC

-177 CGCCGGCCTT TGTATTTGAC GAACTACATA TCCCCCACGT CAGCCTGGAG CCAGAACCCG

-117 TGCCCTATAC TATCTAATCG ACCTTGATCT GGCATATCTA CCAGTTCATG CCGTGGTCGC

-57 TGCGAATTCT TCAGTAATCT GTACCTCTAT ATTATTTTCC TATCCCATAC TCTCACA**ATG**

A.oryzae niiA

-787 TCTGTTTAGG

-777 TACGTACAAT ATTTATGACC TGTGGTGAAA CCTGAGGCAA CAAGGGGGCG CGATTTACCA

-717 GACTGGCGTT CACATACCAA TACAGTGCTT AATTGTAGGT CTCATGGGTG GAATGAGATG

-657 ACCTTCCCTT TCATCTATTC TTAAGAGGAA CAGGGATGGT ACCCACACCA TACCCCGAAG

-597 AGCTCGTGAT GTAATAGACC CTTTCGTAGT ATGCGGGTTT TTATTGAGAT GCCGATATGC

-537 AAACTTGTAG TAAGACTAAT AATAACAGGT GCAATTAATT GAATTTGGGG CCTGTTAAGC

-477 TTTTCGTCCG ACCACCAGTG ACGTCCGTCC GATAAGCCGG GAACGTCGAT AAGGAGGCTA

-417 TCCCGTGCAT CCTGACTAAC CTTTTCCAGT TGTTCCTTAA CTTTGCACCA TCTCCGCGGA

-357 GGAATCCCTG ACCCAATCTT ATGGTGAGGT GCTATCGTCC TTATCAATCC CTGTCACGAT

-297 TGCCGATAAC CCGTCCACAT CCGATCATGG AGATTACCAT TGGGTGTGTC GCTGCATCTG

-237 CTGCGAAACC GCTTCGATGA CCATCAACCA TGAGGTGATT CGATAGTCAA ACATCTCGCC

-177 ACGACCTCAT GGTGTCCACG GAGATGAGAG GCCCGATCCT GGATCTTGGC ACCAGTAGGG

-117 TCTTGCCATA TCGGATGATA GATCATGGGG GACATACTCC TTGATAAATA AAAACGTCCG

-57 CCAACACCGT GCGTTGGGCG TTGAATCTCT TCCACAGTGC TATCCATCCCGGACACA**ATG**

A.oryzae niaD

-897 GACAAATCCT AAGTTAGATG CCGATATATC CGGCACTCTT CAAGCATCAT AGCAATGGCG

-837 TTTTAAATCA CGCAGTTAGG TTGGTGTTCT TCTCATGTGG TAAATCCTCA AGGGTGTAAC

-777 TACAAGTATA CGGTAGACTT CCAAGATTGG CAAAAAAAAG CCAGATCGAG CGATTTTTGC

-717 CTGGATTTAA AGGATCCTGG AGTGCCAATT AAACGTGAGC AATATCCCTC TAACAAACTT

-657 ACCGAGAATT CCTTTGAAGG GTACGTACAG AGTAGTAGCT CCGCTCTAAC AGCCGTGAGC

-597 TCCCATCTGG CCCATTCTCC ACGCGAACCA GGCTCATTCG GACGATAAGC AACAACACTT

-537 CTAATTGGAA GTTGCACATG GCTTCACGGC ATTGTCTCCA TCATTCTTCA AAGAACCAAT

-477 TAATATGACG AAAAAGAAAT CCTGCAAGGT TGAGCGGACA AATGGCTGGA GCCTCGAGAG

-417 TGTTGTGTGG GTCAACGCGA TTTATTTGCC TCAACGCTTG GTTAAGCGGT GGGACGCCGT

-357 CCAGCCTGAA GGCTTGCCCT AATTTCGAGC GTCCCACCTC CCAGAATGAG CTGATTTTGG

-297 GACTCCGCGG AAATCGCTGG TGGGTTCCTA GGAGCATTGC TTATTGTGAC CTTTCTCCGA

-237 GGCGTTATAT GGTAACAAGG AGTTACTGCC GTGACCTAAT AGAATGTCCA CTGCCCGCGC

-177 ATGGGTAACA CGTACTGCGT GCCATCATAC GATAGGCAAG AGTATTTTAG TGTGGTGGAT

-117 TGCCTCCTTA TTTGGAGTTA ATGCACTGGC TGACCATTCG ACTTATTCAT ACTCAGCATC

-57 CTTGTTCATG CTCCCTCCCT CGTAATTTTA CTTTCATTGG CCTTCCCTGC CGGTAAC**ATG**

A.niger niiA

-831 TGTT GAGCTTTTTG CCGCAGTGGG GGCGAGCAGG AAAAGATGGA AGAAAAGAAA

-777 TGGTCCCATC GGCCGACTAC TAGTCAGTCA GGAGTCAGTA GTCAGTGAGT GGACGTCTAG

-717 AACAACTGGT AGTAAGTGAG TAGTAAGCCG TTAGTATGTA GCATGCACCT GCTGCTGGAA

-657 ACCCAACGCA TCCTCAAACC ATTTATCGAA TTCCCTCACA ATGATGTGAC TCCAGCAACT

-597 GCAAGGAGCT CCAGTCTCTT CATTTCCCAC CCGCATCTCT GGTACGTATA CTTCCGATTG

-537 CGGCCAATCA GACGCTTATC AGGAACCGAT AAGGAAGTCG CGATAAGCAA CAGCCTCAAT

-477 GCCTTCGTTT CTCCATCGCA TCTCCGCGGA GTCGTGCTCG CTCGCTCGAT TCATGCAGCC

-417 ACTCCCCAAT TACCGATCGA ACCCATCGCA TTCCCAGTTC TTATCCGTCT TTCCCCCTCT

-357 CCACGAACGG TGGCTCCTCT ATGATCTGCA TGGCGGGTCA AGAGGTCCCT GATTACAGTA

-297 CGATCGAGCA TGGACCCCCC CTCTCTCAAC CGATTACCTC AGAGGTCATA GTCACCCCTC

-237 CAGTCAAGTC CCACGGAGAA ACACAATCCA AACGCCCTAT CGACCGTAAC TCCGCGCAGT

-177 TATCGTTCGC CTCTCGGCGA GATAGATACA ATGATGGTGG TTTGTTGCCT ATGGCTGTGT

-117 AGCGCAACTC CGGTGCCGCT CGGGTTCCCT TTCCCCTGAG GCACAAACAA ACCGCGCTGG

-57 ATTCCTCTTG GCCACAGTCA TCGACGTAGG ATTCCTTGAT CATCCTCTGT CATCACA**ATG**

A.niger niaD

-822 TACTT GCTCAATAAT CCCCGATACC GTGAATACAC GATTAATCAC

-777 GGTCAACTGG AGACTTCGTA ATCCTGCTGC ATTTAATGGG CAATTAATCG GCGGTGGAGC

-717 CCTTTGGAAA CTAAAGTGGA TCTGGCCGCT GTCCTCCTGG GGCTTGTTTT TCCATTACAG

-657 CACAAATTGT CACTAGTATC GAGTTTAGTT TAGTGAGTCT GCTTGCTGCT CGGGAGAGGC

-597 TTGTTTCCAG TTTTTTTGTT CATGCCTGGT TTTTTATACT TTTTTTTTTC GGGACCAGGT

-537 CTCAAGTTCC CACATTCCCG GTAGGAATGT TCGAATGAGA AGCCCGTGTT CTCCCTGTCT

-477 ACCCGACTGG AGGACACACC AGGCCATGAC TAGACCTGCT GGAGAGTACG GAGTATTACT

-417 ATTACATATC TCTGCTTGTT AGTTAACTCA TTGTTCGAAT CCACCGGCTC CCGCGGAAAT

-357 CTCTCCGGCC GAATGTTGGC CCTCACTCAC GCCCTCCCCC CCAAAATAGA TCTCAAACCA

-297 GGTGGAATAC CTCCTGTTTC TTCTCCGTGG ACTTAGTACC AGATCCCGGC CGTTGCTGGT

-237 CACACCCCGT TGAGCTAATC CAAAAGCCGG CAGTTTTGAG GGTTCATCTG TCATCTGGTT

-177 TTTGAGCGAT TAATTTTCCC GTAGCTAACC CTGATCTTCC CCGGATGAGT GTCTTGCTCC

-117 TCTCATCTGC CTTCTGCATG CTCTGTATCG GTCCTAAGCT ATACTCGTCT TCATCCCTGA

-57 TATTGTCTCA TGCCTTATTT GGCACCTCGA GGCCATTGAA CCACATAAAC ACATACC**ATG**

of the *A.nidulans niiA* gene between -138 and -166 nt and -474 and -503 nt (Fig 3.19).

The 5' flanking region of the *A.oryzae niaD* gene has only one copy of TAACTA at -774 nt and one copy of the complementary sequence TCGATA at -550 nt (Fig 3.19). The *A.oryzae niiA* gene has one area at -71 to -99 nt upstream of the ATG in which a similar sequence to TATCTA occurs twice on the complementary strand (Fig 3.19). Within the region -392 to -411 nt upstream of the *A.niger niaD* gene is a copy of TATCTC and the complementary sequence TAGTTA (Fig 3.19). The sequences TATCGA and TATGTA are located separately 5' of the *A.niger niiA* ATG at -628 nt and -677 nt repectively (Fig 3.19).

If these regions are indeed the recognition site of the *areA* encoded protein, the significance born by the eight motifs (Tables 3.8 and 3.9) may be considered to be remote. Despite this, the *nirA* gene product-binding site remains to be identified and functional studies should be conducted to ascertain its sequence and position.

3. 7. 0. HETEROLOGOUS EXPRESSION IN *A.NIDULANS*

DNA mediated transformation allows the investigation of the regulation of structural gene expression by a foriegn regulatory gene. The *nit-2* gene of *N.crassa*, responsible for nitrogen metabolite repression, demonstrated the ability to restore the synthesis of NR in the *A.nidulans areA* loss of function strains, *areA*207 and *areA*19 (Davis and Hynes, 1987). An extension to this study would be to examine the regulation of *A.nidulans* NR, and therefore presumably the *crnA*, *niiA* and *niaD* genes, by the *N.crassa nit-4* gene, accepted as the equivalent of the *A.nidulans nirA* gene.

3. 7. 1. Comparison of the *nit-4* and *nirA* genes at the nucleotide level

The nucleotide similarity between the *nit-4* and *nirA* genes was studied to determine if some sequences were conserved and to investigate the possibility of integration at the homologous site via a DNA-mediated transformation. Integration at the homologous site is more likely to occur when the transforming DNA shares sequence identity with the host chromosomal DNA (Fincham, 1989).

DNA from *N.crassa* and *A.nidulans* wild type strains was digested to completion with the restriction enzymes *EcoR*I, *Xba*I, *Hind*III and *BamH*I. DNA was electrophoresed through a 0.8% agarose gel, transferred to nylon membrane by Southern blotting (2.5.4.a) and hybridised at 50° C, i.e. low stringency, to the hexaprime radio-labelled *EcoR*I/*EcoR*V insert of the *nit-4* plasmid pN4E2A (Fig 2.3). No hybridisation was detected between this probe and the *A.nidulans* DNA (Fig 3.20 ii, lanes A, B, C and D), although as expected, distinct bands are seen for the digested *N.crassa* DNA (Fig 3.20 i, lanes A, B, C and D). The *N.crassa nit-4* gene clearly shares insufficient nucleotide similarity with the *A. nidulans nirA* gene to be detected by Southern blotting.

3. 7. 2. Generation of the *argB*2 *nirA*1 double mutant

It is reasonable to assume that the *nit-4* gene will not integrate at the *nirA* site on transformation, due to inadequate identity between the two genes (3.7.1). In addition, this may result in a low transformation frequency. To combat this it is feasible to perform a co-transformation whereby a gene of substantial identity with the host chromosomal DNA is introduced together with the gene of interest. The *A.nidulans argB* gene, for arginine utilisation, isolated on the plasmid pILJ16 (Fig 2.1), is often utilised in co-transformations to complement an arginine

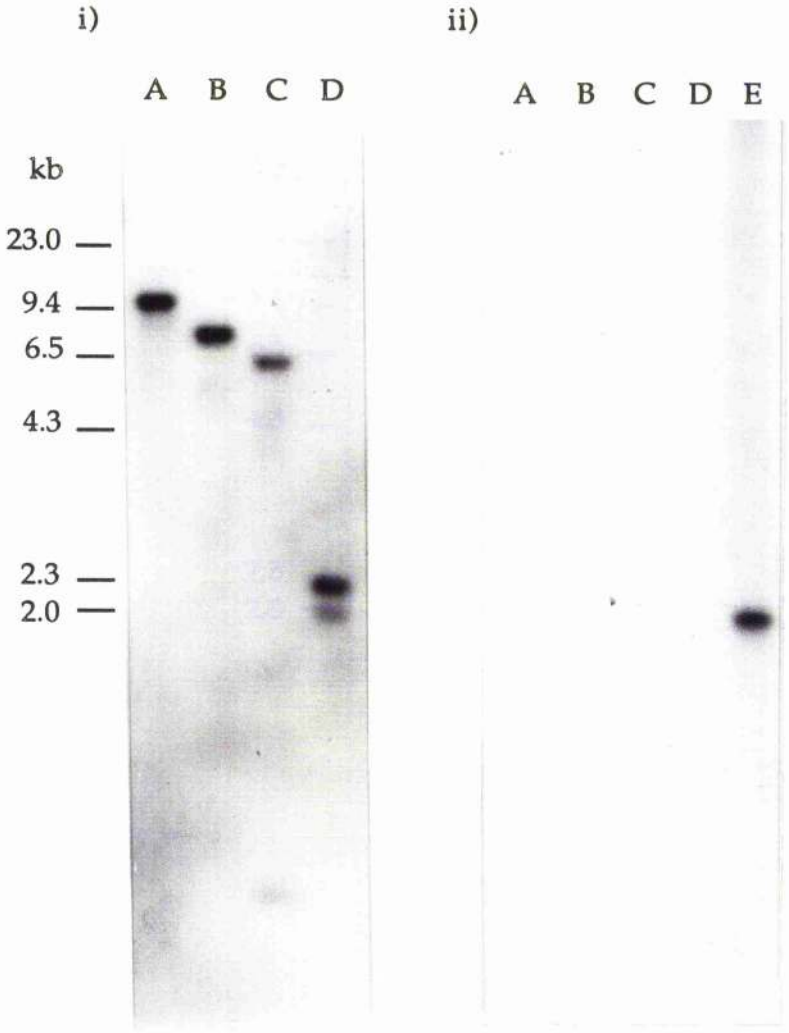


Figure 3.20 Southern blot of, i) *N.crassa* genomic DNA and, ii) *A.nidulans* genomic DNA digested to completion with the restriction enzymes, lane A; *BamH* I, lane B; *EcoR* I, lane C; *Hind* III and lane D; *Xba* I. 10 µg of DNA was electrophoresed on a 0.8% agarose gel, transferred to nylon membrane and hybridised to the hexaprime labelled *EcoR* I/*EcoR* V fragment of pN4E2A. Lane E is 5 ng of *EcoR* I/*EcoR* V pN4E2A fragment.

auxotroph (Johnstone, 1985). The frequency of co-transformation using this system is generally 30%. However, to exploit this system it was necessary to synthesize an *A.nidulans argB2 nirA1* double mutant strain.

The genetic cross of G34 and G834 (2.2.1) yielded a number of progeny of which twenty, from each of four cleistothecia, were tested for growth on MM with glutamate and arginine, or MM with glutamate alone. The number of progeny that appeared to be arginine auxotrophs amounted to eight and were subsequently examined for the inability to synthesize NR when grown on nitrate (this test was considered more accurate than assaying for chlorate resistance when determining the *nirA* phenotype). Only two *argB2 nirA1* double mutants, satisfying the above criterion, were identified. These strains were preserved in silica gel (2.2.3) and named SAA9002 and SAA9003 with the genotypes *biA1 pyroA4 argB2 nirA1* and *yA1 pyroA4 argB2 nirA1*, respectively.

3. 7. 3. DNA-mediated transformation of SAA9003.

A culture of SAA9003 grown in MM with 10 mM ammonium and 10 mM arginine at 37° C, with 300 rpm orbital shaking was harvested after 12 h and protoplasts generated (2.2.2). A co-transformation (2.2.2) with 3 µg of pILJ16 (Fig 2.1) and 10 µg of pNIT-4b (Fig 2.3) was performed. Protoplasts containing the integrated *argB* gene were selected by regeneration on MM with 1.2% sorbitol and 5 mM ammonium.

The seven *argB* transformants were then examined for growth on MM with 5 mM ammonium and MM with 10mM nitrate. Three of the transformants were able to grow on nitrate, implying that the pNIT-4b plasmid had integrated into the host genome and the *nit-4* gene was capable of regulating the expression of the *crnA*, *niaD* and *niiA* genes, restoring the ability to synthesize functional NR. The frequency of co-transformation was calculated as 43%.

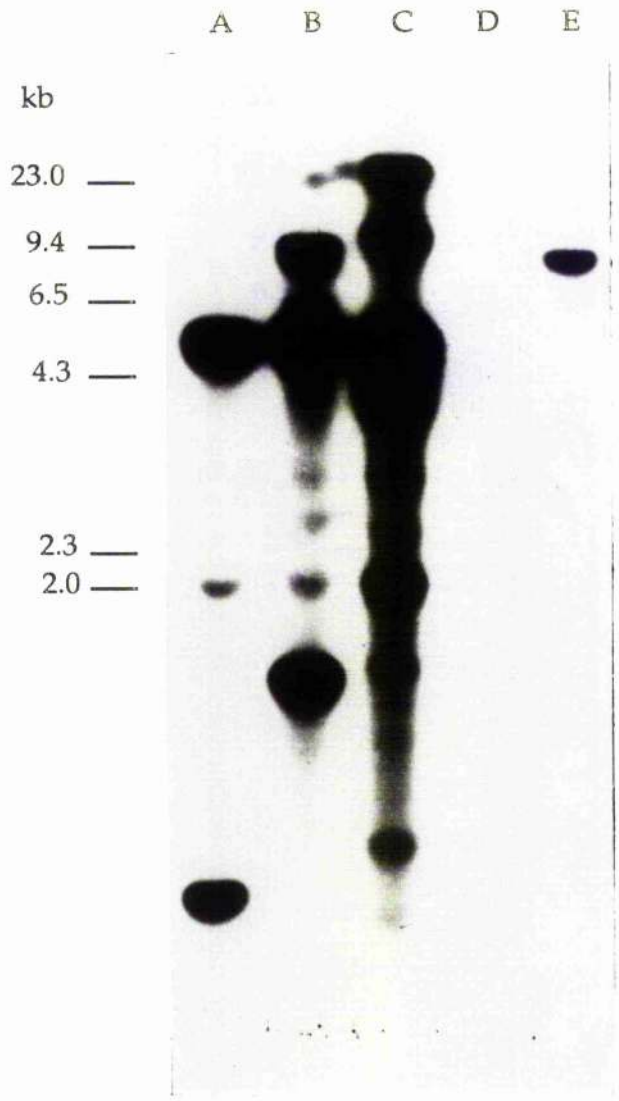


Figure 3.21 Southern blot of genomic DNAs digested to completion with *EcoR* I, transferred to nylon membrane and hybridised to the hexaprime labelled *EcoR* I/*EcoR* V fragment from pN4E2A. Lane A; KHT1, lane B; KHT2, lane C; KHT3, lane D; *A.nidulans* wild type, lane E; *N.crassa* wild type.

To confirm the presence of the *nit-4* gene in the *A.nidulans* transformed strains, 10 μg of genomic DNA extracted from the three transformed strains (KHT1, KHT2 and KHT3), wild type *N.crassa* and wild type *A.nidulans* was digested to completion with *Eco*RI. DNA was electrophoresed through a 0.8% agarose gel alongside *Hind*III cut λ DNA molecular weight markers, transferred to nylon membrane by Southern blotting and hybridised at 64° C to the hexaprime radiolabelled *Eco*RI/*Eco*RV insert of pN4E2A (Fig 2.3). A multiple hybridisation banding pattern for the three transformed strains, KHT1 (Fig 3.21, lane A), KHT2 (lane B) and KHT3 (lane C) is illustrated. The probe hybridises to a single band of 8 kb within the *Eco*RI digested genomic *N.crassa* DNA whereas no hybridisation to the *A.nidulans* wild type DNA is detected, as shown previously (Fig 3.20 ii). The fragment size of 8 kb is larger than expected. The plasmid map indicates that the probe used should hybridise to a fragment of 3.7 kb. A possible explanation for this is that the "*Eco*RI" ends of the genomic DNA in pNIT-4 may not be original and therefore the 3.7 kb band is not a true measure of the *Eco*RI fragment profile in the genome. The measurement of 8 kb is consistent between Southern blots (compare Fig 3.20 i with 3.21) suggesting that the latter explanation is indeed the case rather than there being any fault with the hybridisation.

Multiple integrates are observed for all three transformed strains. It is presumed that the smaller of these fragments may represent incomplete integration of the *nit-4* gene in the proximity of an *Eco*RI site. The larger bands may be as a result of partial integration within a larger *Eco*RI site. Only the bands of approximately 3.7 kb may be a consequence of unaltered integration of the *nit-4* plasmid. Additionally, it is possible that the bands in all three transformed strains represent tandem repeats due to their increased intensity compared with the single band of the wild type.

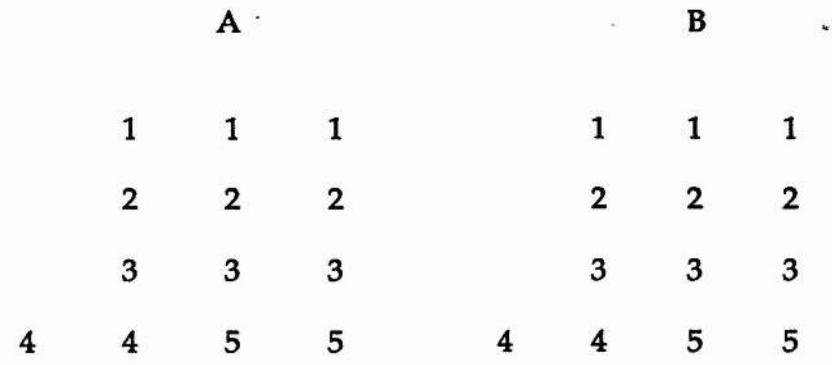
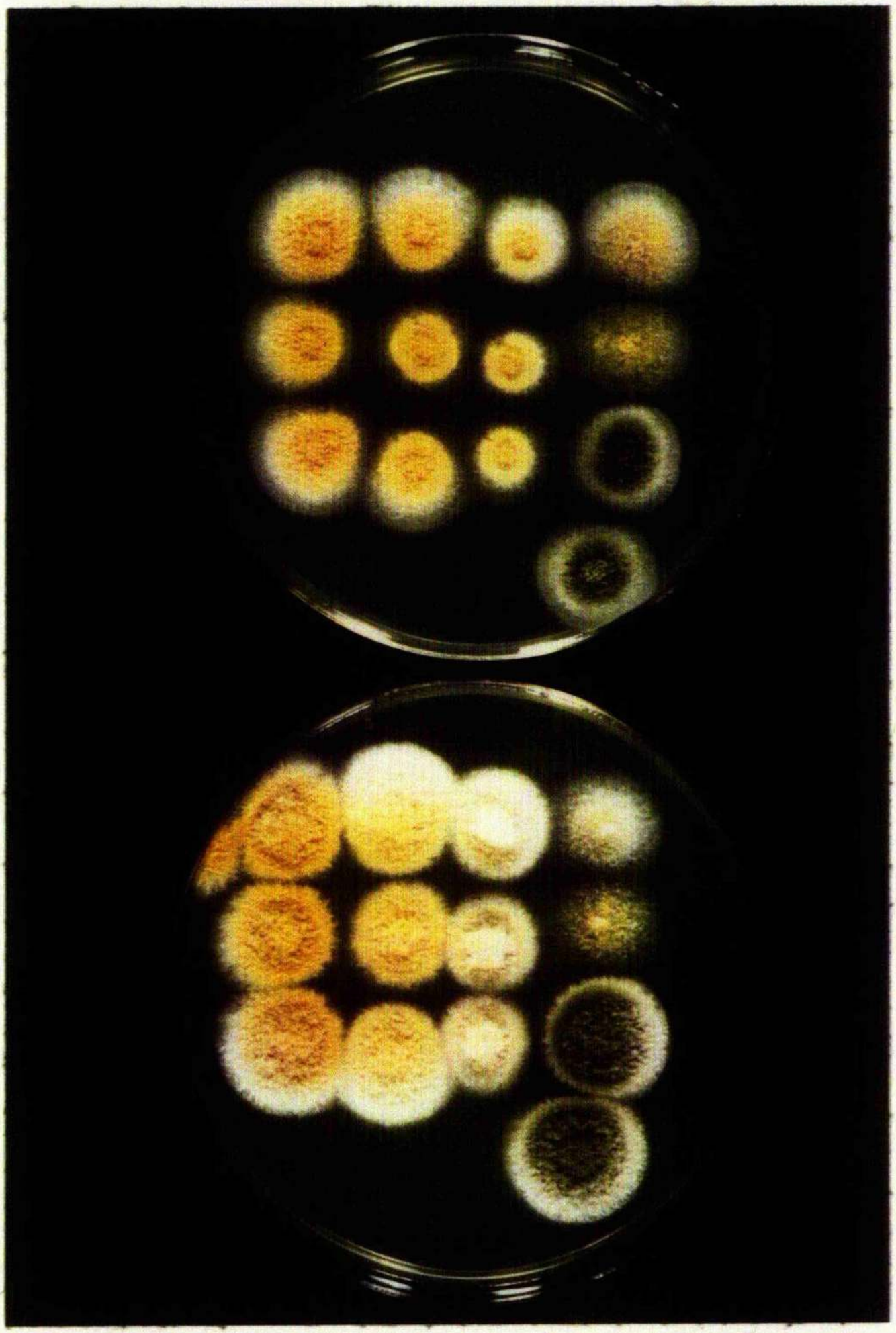


Plate 3.3 Growth characteristics of the KHT strains. From the above key petri dish A is MM with 10 mM nitrate and petri dish B is MM with 5 mM nitrate. The transformed strains shown in triplicate are 1; KHT1, 2; KHT2, 3; KHT3 and the wild type, 4, and *nirA*1, 5, strains are shown in duplicate.



The growth characteristics of the transformed strains were examined in relation to the *A.nidulans* wild type and *nirA*1 strains. The growth rate of the three transformant strains and the wild type differed very little when grown on MM with varying concentrations of nitrate and nitrite (An example of this is illustrated in Plate 3.3). In contrast, the *nirA*1 strain is unable to grow normally on either nitrogen source.

3. 7. 4. Genetic analysis of transformed strains

Genetic crosses between the *A.nidulans* wild type strain and each of the transformed strains allows the identification of the site of plasmid integration. If integration had occurred at the *argB* site the above cross would yield no progeny auxotrophic for arginine. Likewise, integration at the *nirA* site would yield no progeny with the inability to grow on nitrate. However, only progeny prototrophic for arginine can be determined for a *nirA*$^+$ phenotype since with *arg*$^-$ progeny, it is

Table 3.10 The phenotype of the progeny resulting from a genetic cross of the wild type and each of the Kim Hawker Transformant (KHT) strains.

Transformant	Number of randomly isolated progeny	Phenotype of progeny			
		argB$^+$	argB$^-$	nirA$^+$	nirA$^-$
KHT1	153	153	0	112	41
KHT2	193	152	41	88	64
KHT3	180	148	32	124	24

necessary to add arginine to the medium and consequently utilisation of nitrate as sole nitrogen source cannot be phenotypically scored. In contrast, integration occurring at a site other than the *argB* or *nirA* genes would allow the segregation of the *argB*2 and *nirA*1 phenotype in a genetic cross.

Genetic crosses were performed as described (2.1.2) and the number of progeny tested for each cross is shown (Table 3.10). The phenotype of these progeny demonstrates that integration occurred at the *argB* locus on chromosome III in KHT1. Alternatively, integration in KHT2 and KHT3 did not take place at either the *argB* or the *nirA* loci. This supports previous Southern blotting data suggesting that neither plasmid would integrate at the *nirA* locus (3.7.1).

3. 7. 5. Regulation of *A.nidulans* NR by the *N.crassa nit-4* gene

To investigate the regulation of *niaD* gene expression by the *nit-4* gene it was necessary to perform NR assays of the three *nit-4* gene transformed strains grown on varying nitrogen sources.

From the standard curves for NR and protein concentration the specific activity of the NR enzyme in each strain grown in different nitrogen sources can be calculated. The results (Table 3.11) indicate that the NR activities of glutamate grown KHT strains are not significantly different to those of the wild type and are not as great as nitrate grown mycelia, suggesting the *nit-4* gene product does not permit constitutive expression of the *niaD* gene. (The Northern blotting experiments (3.4.4 and 3.5.3) suggest that glutamate is a neutral nitrogen source.) Further evidence for the lack of constitutive *niaD* gene expression is obtained from the NR activities of each strain grown in a second neutral nitrogen

Table 3.11 NR activities, expressed in specific activity (nM NO_2 produced h^{-1} mg $protein^{-1}$), of the transformants; KHT1, KHT2 and KHT3. The wild type and *nirA*1 strains are included for comparison. Standard curves for NR and protein concentration as measured by the method of Bradford (1976) are shown (Fig 2.4).

Strain	Nitrogen Source					
	Ammonium transferred to Nitrate	Urea transferred to Nitrate	Ammonium	Glutamate	Urea	Nitrate and Ammonium
WT	121.0	190.4	1.6	1.7	17.6	6.4
*nirA*1	11.2	<1	1.1	1.6	5.4	2.2
KHT1	223.7	239.6	5.2	13.1	24.8	54.5
KHT2	288.8	281.5	3.3	9.4	23.1	25.9
KHT3	178.3	310.2	4.0	5.9	12.0	22.2

source, urea. Although the NR activities of each transformed strain grown in urea are higher than the corresponding activities for glutamate grown mycelia, they are comparable to that of the wild type grown in urea.

The NR activities of the transformed strains when transferred from either ammonium or urea to nitrate as sole nitrogen source show an increase with respect to the wild type, in addition to the NR activities of the KHT strains grown on glutamate or urea alone. The strain KHT2 demonstrates the highest level of NR when grown initially on ammonium whereas KHT3 indicates greater NR activity when ammonium is replaced with urea as the initial nitrogen source. These observed differences may be a consequence of the data being generated on separate days. Induction of *niaD* gene expression by the *nit-4* gene product is therefore dependent on the presence of nitrate. Increased NR activity of nitrate grown KHT strains, compared with the wild type may be a result of multiple integrates (3.7.3) increasing the *nit-4* gene copy number and possibly the potential to induce the expression of the *niaD* gene. This is particularly plausible since the concentration of the *nirA* gene product is normally at a limiting level (1.4.7).

The NR activity of each transformed strain grown in nitrate and ammonium is higher than the wild type, in addition to the nitrogen metabolite repressed NR levels for each strain, i.e. ammonium grown. This difference may suggest that each KHT strain is slightly derepressed for NR activity. Alternatively, these higher NR levels of the KHT strains, compared with the wild type may be a consequence of multiple *nit-4* gene integrates (3.7.3).

Chapter 4

DISCUSSION

4. 1. 0. INTRODUCTION

The discussion expands on some of the points raised in the RESULTS and will attempt to further explain them. In addition, an outline of ideas for future work is provided.

Initially, the role of the smaller 1.1 kb transcript is discussed followed by a detailed examination of the mechanism of NR autoregulation (4.2.0). The following section (4.3.0.) relates the regulation of the nitrate gene cluster by the *areA* and *nirA* gene products with the conserved motifs within the 5' non-coding regions of the *A.nidulans crnA, niiA* and *niaD* genes in addition to the *A.oryzae* and *A.niger niiA* and *niaD* genes.

An extension of the results regarding the structure of the *crnA* gene and polypeptide is detailed in sections 4.4.0 and 4.5.0, respectively. Section 4.6.0 discusses possible explanations for the observed changes of NR regulation by the *N. crassa nit-4* gene compared with wild type *A.nidulans*.

Finally, proposals for further work to extend our knowledge of the 1.1 kb *crnA* transcript, the regulation of the nitrate gene cluster and the structure of the *crnA* encoded polypeptide are reviewed (4.7.0).

4. 2. 0. REGULATION OF *crnA, niiA* AND *niaD* GENE EXPRESSION

The regulation of *crnA, niiA* and *niaD* gene expression was investigated by studying the mRNA levels of each gene on Northern blots (3.4.0 and 3.5.0) by hybridisation to individual gene-specific probes.

The *crnA* cDNA hybridises to two messages on Northern blots (3.4.0). The larger, inducible message of 1.8 kb, is in agreement with the expected size of the *crnA* mRNA from the sequence analysis. The smaller

1.1 kb message is constitutively expressed and its precise origin not yet determined.

The existence of more than one mode of nitrate uptake has been suggested (Brownlee and Arst, 1983). It was proposed that one system is active in conidia and young mycelia with a second becoming active in older mycelia. Initially, the *crnA* gene was shown to be expressed in young mycelia, i.e. 12 h, (Fig 3.6) leading to the consideration that the smaller message may be transcribed from the second nitrate transport gene of *A.nidulans*, active in older mycelia. However, the abundance of the two transcripts does not alter with respect to each other as the age of the mycelia increases (3.4.5) suggesting that the two transcripts are not from developmentally regulated genes. An alternative explanation is that the mechanism of nitrate uptake in *A.nidulans* is similar to that of barley, in which there are considered to be two systems. One system is constitutive with high affinity and responsible for initial nitrate uptake. Once inside the cell, nitrate is then proposed to induce the synthesis of a second transport protein (reviewed in Wray, 1989). The abundance of the two *crnA* transcripts was not analysed immediately after the addition of nitrate preventing confirmation of this theory. However, the manner in which the two transcripts are expressed, i.e. one constitutive and the other inducible, suggests similarities with the proposed mechanism of nitrate uptake in barley.

That the *crnA* cDNA hybridises equally to the two messages under increasingly stringent conditions would suggest the two transcripts are of very similar sequence or they are transcribed from the same gene. Northern and Southern blots of *crnA* deletion mutants demonstrated that the smaller message is transcribed off the sense or non-sense strand from within the region of the *crnA* gene (3.4.0). The *crnA* anti-sense sequence includes many stop codons decreasing the probability of the presence of a gene. Consequently, it is presumed that the smaller mRNA

is transcribed off the *crnA* gene (3.4.0). This does not lead to the rejection of the possibility of two transport systems as the two transcripts may encode two functional proteins with related but not identical function. The protein encoded by the 1.1 kb message, if functional, would consist of similar membrane-spanning domains as described for the *crnA* encoded polypeptide (Fig 3.5), only fewer. It is possible that this protein may also be capable of nitrate transport but with high affinity thereby allowing initial, low level nitrate uptake when nitrate first becomes available in the immediate environment. Perhaps this leads to the induction of *crnA*, *niiA* and *niaD* gene expression as a result of NR and NIRA dissociation (3.5.4). Similarly, it is possible that the two mRNAs transcribed off the phosphate transport gene, *pho-4*, of *N.crassa* (Mann *et al*, 1989) encode two separate proteins.

Several mechanisms are known to produce transcripts of different function from the same gene but rarely are both synthesized simultaneously (1.3.2). Alternative splicing results in two transcripts from the *niiA* gene (Johnstone *et al*, 1990) but this would not account for a 700 bp difference between *crnA* transcript sizes. A transcriptional termination site within the third intron would result in a transcript of approximately 1.54 kb. Despite this, differential splicing and 5' heterogeneity remain feasible explanations for the existence of the 1.1 kb transcript. In contrast, there is the possibility that the smaller message may be the result of a processing phenomenon of the larger *crnA* transcript or simply an artefact derived from the *crnA* cDNA probe. (Such as a short stretch of DNA that hybridises to a *crnA*-unrelated, constitutively expressed transcript). The latter seems unlikely due to the loss of the 1.1 kb transcript in the deletion mutants, Δ506 and Δ507 (3.4.0) and also in wild type cells grown in glucose limiting conditions (3.4.2).

The expression of the *crnA* gene is similar to the *niiA* and *niaD* genes in that it is inducible with nitrate and nitrite (3.4.1). When mycelia

synthesizing the *crnA* mRNA are introduced into an unfavourable environment, synthesis is very rapidly arrested. Northern blots demonstrate that when mycelia are exposed to an environment lacking glucose, transcription of the *crnA* gene ceases and all remaining mRNA is degraded within 15 min (3.4.2) . These results suggest that expression of the *crnA* gene is tightly regulated (Shapiro *et al*, 1988), the mRNA having a half life of less than 7.5 min. This is also true of many other filamentous fungal mRNAs. For example, the *A.nidulans* arginase gene transcript has a half life of 2.5 min (Cybis and Weglenski, 1972). An explanation for the tight regulation is that these genes offer a rate limiting step in a metabolic pathway and need to turn over rapidly in response to changing nutritional requirements (Goldberg and Dice, 1974).

Functional *nirA* and *areA* genes are shown to be required for *crnA* gene expression (3.4.3) and similarly *niiA* and *niaD* gene expression (3.5.2). In addition, from experiments presented here, NR would appear to play a role in the regulation of the expression of the nitrate gene cluster as originally suggested by Pateman and Cove (1969). That several *niaD* and one *cnxE* mutant strains result in constitutive expression of the *crnA*, *niiA* and *niaD* genes (3.4.4) would suggest that the NR holoenzyme, rather than the *niaD* encoded polypeptide, is normally required for the repression of the three genes. This is in agreement with Pateman and Cove's theory that NR may interact with the *nirA* gene product as a mechanism to prevent expression of the *niiA* and *niaD* genes (1.4.7). The results presented in this thesis (3.4.0) suggest that the *crnA* gene should be included in this model.

Previous theories (Cove 1979) have proposed that the *nirA* gene product is required for the expression of the *niiA* and *niaD* genes and is only active in the presence of nitrate. The results presented here support the view that the *nirA* gene product is required for *niiA* and *niaD* gene expression (3.5.2) and demonstrate that *crnA* gene expression is also

dependent on a functional *nirA* gene product (3.4.3). It follows that the *nirA* gene may, itself, be constitutively expressed since certain mutations within the *niaD* and *cnxE* genes lead to constitutive synthesis of the *crnA*, *niiA* and *niaD* mRNAs. Certainly the *nirA* gene product has a low copy number, consistent with the notion of a basal level of constitutive transcription.

If the *nirA* gene is not constitutively expressed then an alternative mechanism must be operating to allow the constitutive expression of the *crnA*, *niiA* and *niaD* genes in certain *niaD* and *cnxE* mutant backgrounds. It is possible that this may occur as a result of local changes in the chromatin structure (1.3.1) (Weisbrod, 1982). Already discussed is the fact that transcriptional activation of genes is preceded by changes in the chromatin structure (1.2.0.) This is promoted by a conformational change in the DNA helix, from right-handed, B-DNA to left-handed, Z-DNA. Such a change is stabilised or perhaps results from DNA supercoiling, binding of specific proteins or ions and methylation of DNA (Nordheim *et al*, 1981). Undermethylation of all, or a sub-set of CG dinucleotide sites within a given gene is proposed to alter the chromatin structure (Groudine *et al*, 1981). In addition, methylation of DNA is assumed to hold nucleosomes in position, control regions of active genes being wound onto nucleosomes (Kolata, 1985). Various mutations within the *niaD* gene may result in loss of CG residues and therefore, possibly demethylation of that region of the chromosome, allowing the formation of Z-DNA, sliding of nucleosomes and transcriptional activation.

Strains with mutations including all or part of the upstream regions of the *niaD* gene, i.e. *niaD*5 and Δ509 do not express the *niaD* gene (3.5.4), as expected, but do allow constitutive expression of the *crnA* gene (3.4.4). It would appear then, that functional NR is not necessary for the induction of *crnA* gene expression but is required for the repression of *crnA* mRNA synthesis and presumably also for the repression of *niiA*

and *niaD* gene expression when nitrate is not available. However, the alternative explanation described above is valid here also.

One interpretation of these results supports Pateman and Cove's model for the interaction of NIRA with the NR enzyme and provides evidence for their involvement in the regulation of *crnA* gene expression. The *nirA* gene may be constitutively expressed and under non-inducing conditions its product may bind to the NR enzyme. When nitrate becomes available the NR-NIRA dimer dissociates allowing the *nirA* encoded protein to bind to *cis*-acting regulatory sequences 5' to the *crnA*, *niiA* and *niaD* genes, inducing their expression. When the *nirA* binding site of NR is non-functional, interaction with the *nirA* gene product is prevented and constitutive expression of the *crnA*, *niiA* and *niaD* genes ensues. This model shows similarities to that of the GAL4/GAL80 regulatory system in *S.cerevisiae* (1.3.1). The following comparison between the two systems is entirely speculative and not based on the structure of the NIRA polypeptide since this information is unavailable. The *nirA* gene product may behave in a similar manner to GAL4, its N-terminus possibly interacting with *cis*-acting regulatory seqences of the *crnA*, *niiA* and *niaD* genes. When nitrate is present the C-terminus of NIRA may interact with transcription factors, bound to the promoter sequences of the nitrate gene cluster. Therefore, NR appears to have a similar function to GAL80, competing with the transcription factors, for NIRA. When nitrate is absent, evidence suggests that NR is capable of interacting with NIRA, bringing about the repression of *crnA*, *niiA* and *niaD* gene expression. Thus, mutations in the *nirA* gene resulting in the loss of the C-terminus and mutations of the *niaD* or *cnx* genes resulting in an altered NR would allow constitutive expression of the nitrate gene cluster. However, there is the question of how nitrate enters the cell initially. Perhaps the 1.1 kb *crnA* transcript does encode a constitutive, high affinity nitrate transporter. Further, the residual NR

may not be in large enough supply to bind with the *nirA* encoded protein in order to prevent expression of the structural genes. This consideration may be of little consequence as the *nirA* gene product has been shown to be present in near limiting concentrations (1.4.7) (Cove,1979)

A second model is for the NR enzyme to interact as a *trans*-acting negative regulatory protein with *cis*-acting upstream repressor sequences of the *nirA* gene. In this way free NR (not complexed with nitrate) may prevent *nirA* encoded polypeptide synthesis, which in turn prevents the expression of the *crnA*, *niiA* and *niaD* genes. When the DNA binding site of NR is disrupted it is unable to interact with upstream sequences of the *nirA* gene. This results in the constitutive expression of the *nirA* gene and in turn the nitrate gene cluster.

Either model requires the NR enzyme to have a motif (Fig 1.3) allowing the interaction with either another protein or DNA. No such motif (Mitchell and Tjian, 1989) could be identified within the *niaD* encoded polypeptide although it is possible that these motifs vary greatly. It is perhaps worth noting that the AA sequence of the *niaD* gene is rich in acidic residues; glutamic acid and aspartic acid together contributing to 12% of the polypeptide. Proline, another AA associated with DNA-binding (1.3.1) constitutes 6% of the *niaD* encoded protein. However, a similar analysis of the *niiA* encoded polypeptide reveals 13% acidic residues and 4% proline residues. The lack of any significant difference, in this respect, between the *niiA* and *niaD* encoded polypeptides may suggest that the distribution of such residues is more important than the overall quantity.

Finally, since both the *areA* and *nirA* gene products are required for the expression of the *crnA*, *niiA* and *niaD* genes it is considered that the *areA* gene is expressed constitutively in addition to the *nirA* gene.

Finally, since both the *areA* and *nirA* gene products are required for the expression of the *crnA*, *niiA* and *niaD* genes it is considered that the *areA* gene is expressed constitutively in addition to the *nirA* gene.

4. 3. 0. UPSTREAM NUCLEOTIDE SEQUENCE ANALYSIS

Both the *areA* and *nirA* gene products have been shown to be responsible for the regulation of *crnA*, *niiA* and *niaD* gene expression (3.4.3 and 3.5.2). Although the NR enzyme also appears to play a role in regulation, it is believed that its effect is exerted by dimerisation with the *nirA* encoded protein. In contrast, it is generally agreed that at least one, and possibly both of the *areA* and *nirA* gene products interact directly with the upstream regions of the *crnA*, *niiA* and *niaD* genes (1.4.7). It was proposed that regulation of the nitrate assimilatory structural genes in *A.nidulans* is similar to *A.oryzae* and *A.niger*. Support for the proposal was gained from recent evidence indicating that the regulatory genes *nit-2* and *nit-4* of *N.crassa* are able to substitute for the *areA* and *nirA* genes of *A.nidulans* respectively. In an attempt to identify putative *cis*-acting regulatory sequences involved with *areA* and *nirA* binding the 5' non-coding sequences of the *A.nidulans niiA* and *niaD* genes were compared. Similar comparisons between genes of related function have identified motifs that were later shown to be involved with regulatory gene product binding by DNase I footprinting, *lacZ* fusions and protein-binding assays (1.2.1).

Eight motifs common to the upstream regions of the *A.nidulans niaD* and *niiA* genes were subsequently compared with the upstream regions of the *crnA* gene (3.3.2) and the *A.oryzae* and *A.niger niiA* and *niaD* genes (3.6.2). Only motif 5 is conserved to any degree and common to all seven genes. The significance of this is unclear although one

possibility is that the motif represents a *cis*-acting sequence recognised by either the *nirA* or *areA* encoded proteins. Upstream sequence comparisons were therefore relatively unsuccessful in identifying homologous sequences for this system. Certainly, no substantial nucleotide identity is observed, unlike the 50 nt sequence common to both the *A.nidulans* and *A.niger gpdA* genes (Punt *et al*, 1990).

Three different models for the way in which the *areA* and *nirA* encoded proteins interact to bring about the expression of the *niiA* and *niaD* genes were suggested by Scazzocchio and Arst (1989) (1.4.7). The presence of one *cis*-acting sequence, i.e. motif 5, is probably all that is required for the activation of the *niiA* and *niaD* genes in two of these models. First, the *nirA* and *areA* encoded proteins may dimerise before interacting with upstream sequences of the *niiA* and *niaD* genes. Second, the *areA* encoded protein may be required for the expression of the *nirA* gene. In this case only the NIRA protein interacts with upstream sequences of the *niiA* and *niaD* genes. The sequence of both the *areA* gene (Kudla *et al*, 1990) and the *nirA* gene (Scazzocchio, unpublished data, reviewed in Crawford and Campbell, 1990) indicate the presence of a zinc finger, involved with DNA-binding (1.3.1). This information suggests that both regulatory proteins interact directly with upstream regulatory sequences of genes, supporting the "cascade" model and also a third model in which the two proteins interact independently with the *cis*-acting sequences of the *niiA* and *niaD* genes. Results presented in this thesis suggest that the *crnA* gene should be included, along with the *niiA* and *niaD* genes, in these models.

Recent evidence suggests that four areas containing a core sequence of consensus TATCTA, upstream of the *niaD* gene are recognised by the *nit*-2 encoded protein (Fu and Marzluf, 1990). Similarities between the zinc finger DNA-binding domain of the *nit*-2 (Fu and Marzluf, 1990) and

areA (Kudla *et al*, 1990) gene products imply that the AREA protein will interact with identical *cis*-acting elements. Computer-aided searches identify several TATCTA or related sequences within the *niiA-niaD* intergenic regions of *A.nidulans*, *A.oryzae* and *A.niger* (3.6.2) in addition to the 5' non-coding sequences of the *crnA* gene (3.3.2). All of the identified elements may be necessary for AREA-binding although their importance should be examined by protein-binding assays (4.6.2). The existence of motif 5 may be coincidental or serve an entirely separate purpose. One explanation is that it is the recognition site for the NIRA protein.

4. 4. 0. STRUCTURAL ANALYSIS OF THE *crnA* GENE

The results of the sequencing and primer extension analysis have enabled an investigation of the *crnA* gene at the nucleotide level and comparison with other filamentous fungal genes.

Two TATA motifs are situated 19 and 51 bp upstream of the transcriptional start site of the *crnA* gene (3.3.2) and are considered to be constituents of the core promoter sequence. TATA boxes are thought to be of specific importance in those genes requiring a high level of expression (1.2.1). This is in agreement with the results of the Northern blotting studies, where the *crnA* gene is shown to be inducible, and degraded very rapidly, i.e. under tight regulation (3.4.0). It is unusual in *A.nidulans* to find two TATA motifs although the phenomenon is observed in the sesquiterpene cyclase gene of *F.sporotrichioides* (1.2.1).

The TC block, generally found upstream of the transcriptional start site of many filamentous fungal genes is downstream of the *crnA* gene transcriptional start site. This raises the possibility of multiple initiation

sites. However this is not detected on Northern blots and would not account for the 700 bp difference between the two *crnA* transcripts.

The 190 bp leader sequence is long compared with many *A.nidulans* genes (1.2.1) and would support the notion that a second initiation site does exist further downstream, especially since filamentous fungal genes generally have multiple initiation sites (1.2.1). Despite this, primer extension studies imply that this leader is dominant to any other that may exist and, although more unusual in *A.nidulans*, several *N.crassa* genes have leader sequences as great as 400 nt for example the *qa-4* gene of *N. crassa*, for quinic acid utilisation (Rutledge, 1984).

The coding region of the *crnA* gene is interrupted by three introns, two near the 5' end and the third close to the 3' end. This pattern is in general agreement with other *A.nidulans* genes showing conservation of intron position wihin a species. In addition, it has been shown that to a certain degree intron position can be related between similar genes of different organisms. This may demonstrate phylogenetic relationships between species.

The consensus AATAAA, or any other polyadenylation site (1.2.1) is not present in the *crnA* gene although the long TG-rich sequence may exert an effect on poly(A) addition. The YGTGTTYY consensus, possibly involved with termination of the transcript, can not be identified either (1.2.1). Few fungal genes are known to posess the YGTGTTYY motif rendering its significance unclear.

4. 5. 0. STRUCTURAL ANALYSIS OF THE CRNA POLYPEPTIDE

The putative structure of the *crnA* encoded protein was calculated using the method of Eisenberg *et al* (1984) (1.5.2.a). It is similar to other

membrane proteins in that apolar AA residues are arranged in groups proposed to span the membrane. These hydrophobic regions are connected by hydrophilic loops on the external and internal surfaces. The models of several transport proteins show one relatively long hydrophilic region within the polypeptide, for example the phosphate transporter of *N.crassa* (Mann *et al*, 1989), the quinic acid permease of *N.crassa* (Geever *et al*, 1989) and the lactose carrier of *E.coli* (Kaback *et al*, 1990). This feature is present in the CRNA protein but its functional significance, if any, is unclear. In many cases the hydrophilic region is situated centrally, surrounded by an equal number of hydrophobic domains, implying that transport proteins are a result of an internal gene duplication in an ancestral transporter (Maiden *et al*, 1987). In contrast, the *crnA* encoded protein is asymmetrical and it lacks motifs common to both sides of the hydrophilic region suggesting that the *crnA* gene evolved by means other than duplication. Alternatively, the loss of two hydrophobic domains may have occurred following duplication due to their presence having little purpose within the protein.

Eisenberg's method is calculated on the proviso that all membrane domains require 21 AAs (1.5.2.a). It is possible for shorter domains to reside within the membrane provided they are surrounded by domains of 21 AAs for stability. Therefore there are limitations with Eisenberg's method and the protein may not assume quite the predicted structure. In addition, domains 1 and 8 may be surface seeking (Table 3.4), according to Eisenberg, and would therefore result in an altered prediction of the CRNA protein structure. It should also be noted that the protein would have a 3D configuration within the membrane. A prediction of this structure is impossible given the information available.

That no ATP binding site (1.5.2.b) is evident would suggest that nitrate transport by the CRNA protein is not linked directly to ATP

utilisation. Perhaps this is not surprising since all such precedents are Gram-negative bacterial periplasmic transport systems, for example histidine transport in *S.typhimurium*. (1.5.1). However, a human calcium transport protein (Riordan *et al*, 1989) has a very different ATP-binding site to the one identified by Walker (1982) and referred to in this study. There is the possibility then that the CRNA protein does have ATP binding activity but the AA motif determining this is as yet unidentified.

No N-glycosylation motif has been identified within the *crnA* AA sequence. This is particularly unusual as many, although not all, transport proteins are N-glycosylated in at least one position. Most proteins synthesized on the rough ER are glycoproteins being carried, via the Golgi apparatus, to the cell membrane (Lodish, 1988) where they become inserted with the oligosaccharide generally on the exterior of the cell. That CRNA is not glycosylated does not exclude the possibility that it is synthesized on the rough ER. However, other methods for the insertion of proteins into membranes have been proposed. (1.5.3) It has been considered that proteins may be synthesized by ribosomes attached to the plasma membrane, insertion occurring via signal sequences and stop-transfer sequences located within the protein (1.5.3). Alternatively, the protein may be synthesized within the cytosol and subsequently diffuse into the more energetically favourable environment of the membrane by the membrane-trigger or helical-hairpin hypotheses (1.5.3.).

4. 6. 0. THE EFFECT EXERTED BY THE *nit-4* GENE ON *A.NIDULANS* NR ACTIVITY

The *N.crassa nit-4* gene was introduced into an *A.nidulans* strain with a non-functional *nirA* gene via a DNA-mediated transformation system. That none of the transformants showed integration at the

homologous site is not surprising since the *nit-4* and *nirA* genes were shown, by Southern blotting, to share little similarity at the nucleotide level (3.7.1). Transformation of *A.nidulans* is similar to *S.cerevisiae* in which a number of integration events can arise (Hicks *et al*, 1979): type I integration, where a single cross-over at a site of homology produces a direct repeat; type II in which a single cross-over occurs at another site and type III integration where a double cross-over results in a gene conversion event. It is likely that integration in the transformed strain, KHT1 is type I, occurring at the *argB* locus whereas the strains KHT2 and KHT3 are a result of type II integration.

The transformed strains were all capable of growth on nitrate, as sole nitrogen source, implying the *nit-4* gene does not alter wild type NR synthesis in *A.nidulans*. Additionally, it is plausible that the *nit-4* gene product recognises the *cis*-acting upstream sequences of the *crnA*, *niiA* and *niaD* genes with which the *nirA* gene product normally interacts.

The results of the enzyme assays indicate that NR activity in the KHT strains is generally twice that of the wild type (3.7.4). This effect may be a consequence of the mutiple *nit-4* integrates, a feature of I and II integration types, and that in wild type *A.nidulans* the *nirA* gene product is normally limiting for NR synthesis (Cove, 1979).

The *nit-4* gene does not permit constitutive expression of *A.nidulans* NR under normal conditions, i.e. when no other genes are altered. If models regarding NR and NIRA interaction presented elsewhere in this thesis (3.5.4) are correct it would suggest that *A.nidulans* NR is able to interact with either the *nit-4* gene product or the upstream sequences of the *nit-4* gene when nitrate is absent.

Slight derepressed NR activity is apparent when the *nit-4* transformants are grown in nitrate and ammonium. However, this effect is not total (only 25% of the NR activity produced by the wild type when

grown on nitrate as sole nitrogen source) and therefore is proposed to be a consequence of the multiple *nit-4* gene integrates, allowing increased NR activity in the presence of nitrate.

Derepressed *nirA* alleles have been identified and considered to be a result of loss of interaction between the AREA and NIRA proteins (Tollervey and Arst, 1981) however there is no experimental evidence to suggest this actually occurs. An alternative explanation for the apparent derepression of NR activity in the transformed strains is that the partial *nit-4* gene integrates (3.7.3) synthesize incomplete proteins which are unable to interact with the AREA protein. According to this explanation, derepression is incomplete due to the observation that not all *nit-4* integrates are partial (3.7.3). Furthermore, a complete *nit-4* gene product may have a different structure to the *nirA* encoded protein rendering dimerisation with the *areA* gene product impossible. It is interesting that in a similar experiment involving the *nit-2* and the *areA* genes, *A.nidulans* NR was regulated normally apart from slight derepressed activity (Davis and Hynes, 1987).

These results suggest that AA motifs necessary for protein dimerisation and DNA-binding within the *nit-4* and *nirA* encoded proteins have been conserved during evolution despite the difference in codon usage. That the function of the two proteins is identical would indicate related regulatory mechanisms for nitrate assimilation between *N.crassa* and *A.nidulans*.

4.7.0. PROSPECTS

4.7.1. Analysis of the 1.1 kb *crnA* transcript

It would be interesting to extend our knowledge of the 1.1 kb message to which the *crnA* cDNA hybridises. Provided this message is

transcribed off the *crnA* gene it may be possible to discover its position by using smaller restriction fragments from pSTA4 (Fig 2.2) as probes against RNA on Northern blots. This procedure was used by Mann *et al* (1989) to examine the difference between the two messages transcribed off the *pho-4*, phosphate transport gene of *N.crassa*.

4. 7. 2. Regulation of *crnA, niiA* and *niaD* gene expression

To further investigate the way in which *crnA, niiA* and *niaD* gene expression is regulated by the *areA* and *nirA* gene products and to identify *cis*-acting regions of DNA with which they may interact, protein-binding studies could be conducted, viz gel-shift assays and DNase I footprinting.

Part of the *nit-2* gene product including the zinc finger domain, over expressed in *E.coli*, has been shown to recognise 5' non-coding regions containing several copies of the sequence TATCTA of the *nit-3* and *niaD* genes of *N.crassa* and *A.nidulans*, respectively (3.6.2). This raises a question regarding the significance of motifs 1-8 (3.6.1). However, it is anticipated that if any of these motifs are *cis*-acting regulatory receptor sites for the NIRA or AREA proteins differences in their migration will be observed when used in protein-binding experiments against protein extracted from wild type mycelia grown in various nitrogen sources and similarly from strains mutant for the *nirA* and *areA* regulatory genes, viz *nirA*1, *nirA*^{c}1, *areA*19 and *xprD*1.

A more sensitive method of protein-binding analysis would be to use purified *nirA* and *areA* gene products rather than fractionated cellular protein. The cloning of both genes into high-copy expression vectors allows the synthesis of large quantities of the protein encoded by each gene. Over expression of the *N.crassa qa-1F* gene product (the activator regulatory gene for the quinic acid assimilation pathway) in an

insect (*Spodoptera frugiperda*) cell culture using a recombinant baculovirus was found to be more successful than over expression in *E. coli* (Baum *et al*, 1987). In particular, baculovirus permits synthesis of the whole protein from a single clone whereas over expression of large proteins in *E.coli* is impossible, requiring the use of subclones for synthesis of protein fragments.

Plasmid vectors containing reporter genes, such as the pAN923 (van Gorcom *et al*, 1986) series of vectors, with the coding sequences of the *lacZ* gene of *E.coli*, can be used to determine which upstream sequences are required for the expression of the *crnA*, *niiA* and *niaD* genes. Fragments of varying lengths from upstream regions of all three genes can be ligated separately into the multiple cloning site of the pAN923 vectors in all three reading frames. This results in the introduction of the covalently linked fragment immediately upstream of the *lacZ* coding sequences. Any promoter or *cis*-element harnessed in this position will drive the expression of the *lacZ* gene under inducing conditions. *A.nidulans* wild type transformed with the vector and grown under inducing conditions for *crnA*, *niiA* and *niaD* gene expression will produce the enzyme ß-galactosidase provided the above criterion is satisfied. The chromogenic chemical, X-gal (2.4.1) can be used to determine the quantity of ß-galactosidase produced in the fungal cellular extracts. A similar approach was used to investigate the sequences responsible for the regulation of the *S.cerevisiae* allantoate permease (Rai *et al*, 1989).

Recently, a vector, pTRAN2-1b, was constructed in which the *A.nidulans niiA-niaD* intergenic region was ligated between the coding sequences of the *E.coli lacZ* and *uidA* (encoding ß-glucoronidase) genes (Grieves *et al*, 1990). The *niiA* promoter was linked to the *uidA* gene and the *niaD* promoter to the *lacZ* gene. The plasmid may be transformed

into various species and strains and the regulation of the *niiA* and *niaD* genes examined by measuring the activity of the ß-glucoronidase and ß-galactosidase enzymes with the chromogenic substances 5-bromo-4-chloro-3-indolyo glucuronide (X-gluc) and X-gal (2.4.1), respectively. In particular it would be interesting to determine the regulation of *A.nidulans niiA* and *niaD* gene expression in *A.oryzae* and *A.niger*.

The role played by the NR enzyme in the regulation of *crnA*, *niiA* and *niaD* gene expression could perhaps be further investigated by site-directed oligonucleotide-directed mutagenesis at various positions within the *niaD* gene. This method has advantages over other types of mutagenesis, viz transposon, random and localised random, in that a single pre-determined base may be altered (Botstein and Shortle, 1985). The desired mutation is synthesized *de novo* as a short oligonucleotide which is then incorporated into a longer segment of DNA. Although various methods have been developed, the simplest way this can be done is to anneal the oligomer to the single-stranded form of the wild type gene. The oligonucleotide then serves as a primer for extension by DNA polymerase with the wild type sequence as the template.

The major problem associated with this type of mutagenesis is that extension by DNA polymerase is not totally accurate, requiring the determination of the nucleotide sequence of the mutated gene to ensure the only mutation is that introduced via the oligomer (Botstein and Shortle, 1985). The precise arrangement of the AAs consituting a protein determine its function. Therefore, the mutagenesis of specific AAs may provide information as to the role played by that AA. Certain mutations of the *niaD* gene may affect the NR protein structure in such a way that it is no longer able to function as a regulatory molecule.

4. 7. 3. Structure of the *crnA* encoded protein

A full investigation of the structure assumed by the CRNA protein and the specific AAs required for the transport of nitrate is necessary for a complete understanding of nitrate uptake. It would be possible to obtain the purified CRNA protein by a similar method described for the *nirA* and *areA* proteins (4.5.2), although a full length cDNA would have to be obtained. X-ray crystallography may reveal the 3-D structure of the purified protein, although this method is laborious and complex (Eisenberg and Hill, 1989). A further problem is that the structure may be altered due the protein being removed from its native environment of the membrane.

Several methods have been devised to determine the orientation of *E.coli* trans-membrane proteins and to confirm the location of the trans-membrane domains. The orientation of the lactose carrier, *lacY*, of *E.coli* was investigated by (MacIntyre *et al*, 1989) fusing the transcription activator sequences and varying lengths of the *lacY* gene to the *ompA* gene, encoding the *E.coli* outer membrane protein. Transformation of an *E.coli* strain with a particular construct allows the LACY-OMPAfusions to be transported to the plasma membrane where they are incorporated. All OMPA exposed at the periplasmic surface is accessible to trypsin, upon plasmolysis of the bacterial culture, and can be immunoprecipitated. The orientation of the *crnA* gene product could be investigated in a similar manner provided CRNA is capable of incorporation into the *E.coli* plasma membrane. Immunoprecipitation of OMPA is dependent upon after which hydrophobic membrane-spanning domain of a particular membrane protein it is fused. If the model for the structure of the *crnA* encoded protein is correct it is expected that OMPA would be immunoprecipitated when fused after the CRNA hydrophobic domains 1, 3, 5, 7 and 9.

A similar method, using the alkaline phosphatase gene of *E.coli* was devised by Manoil and Beckwith (1986). Recent experiments, using this technique have established a model for the structure of the *E.coli* lactose permease (Kaback *et al*, 1990). Initially, there were two models; one devised from the method of Kyte and Doolittle (1982) indicated twelve membrane-spanning domains (Foster *et al*, 1983). In contrast, spectroscopy had revealed the presence of fourteen membrane-spanning domains (Vogel *et al*, 1985). Interestingly, the results of the alkaline phosphatase fusions confirm the former to be correct. Varying lengths of the *crnA* cDNA can be attached to the alkaline phosphatase gene and ligated into an expression vector, as before. After transformation, media containing a culture of transformed cells is assayed for alkaline phosphatase activity. The enzyme is only active when exported to the cell surface resulting in differential activity depending on after which hydrophobic domain the alkaline phosphatase gene is fused. Again, alkaline phosphatase activity would be expected when the gene is fused after CRNA hydrophobic domains 1, 3, 5, 7 and 9. However, this method is also dependent on the ability of the *crnA* encoded protein to be incorporated into the *E.coli* plasma membrane.

To discover which AAs are involved directly with nitrate transport it would be necessary to perform site-directed mutagenesis with the *crnA* gene. Active regions of transport proteins have successfully been identified by this method (Cox *et al*, 1988). Specific AA residues should be chosen for alteration, particularly charged residues and cysteine residues within membrane-spanning domains. It is energetically unfavourable to insert polar AAs within membranes and therefore it would be assumed that they must serve a function. Cysteine residues are often associated with a specific function, due to their ability to form disulphide bonds. In addition, a particular histidine residue within the *E.coli* lactose carrier has

been identified to play a role in transport (Kaback *et al*, 1990). However, the *crnA* encoded polypeptide contains no such membrane-buried histidine residues. Other "functional" AAs may include the various membrane-buried proline residues and other residues that are considered to break the α-helical structure within the membrane, i.e. glycine and serine (3.3.4). The AAs: R-87, C-90, C-140, C-144, C-219, C-325, C-367, R-368, D-410 and E-429 should be chosen for mutagenesis, initially. It is important to ensure that the alteration does not result in an AA that could serve a similar function. A change to a hydrophobic AA would be preferable. Site-directed mutagenesis could be performed as described previously and all mutants sequenced.

The mutated *crnA* gene could be covalently linked into pILJ16 housing the *A.nidulans argB* gene (Fig 2.1) and transformed into an *A.nidulans argB*2 *crnA*1 double mutant. Transformants can be selected for on the basis of *argB* complementation. To assess which *crnA* mutants have altered nitrate transport the transformants should be tested for growth on MM with 5 mM nitrate and 20 mM caesium chloride. The *crnA*1 strain grows poorly on such media although the wild type grows normally (Plate 3.1). Therefore, transformants capable of growth on caesium chloride would be presumed to be unaltered for nitrate transport. If the converse was true and growth is not sustained it is likely that nitrate transport had been disrupted by the mutation.

A further prospect is to investigate the possibility of *A.nidulans crnA*, *niiA* and *niaD* gene transformation into the genome of higher plants, with the aim of increasing the rate of nitrate assimilation and hence, crop yield. In particular, *crnA* gene transformation would be interesting as no nitrate transport gene of crop plants has as yet been isolated.

For dicotyledonous plants transformation of foreign genes has become more or less routine provided regeneration from tissue explants or cultured cells is possible. Such transformation systems exploit the transfer function of the Ti plasmid from *Agrobacterium tumefaciens*, whilst eliminating its tumerogenic capabilities. Successful transformation systems using this method have been developed in Tobacco (Barton *et al*, 1987), Rape (Pua *et al*, 1987) and Cotton (Firoozabady *et al*, 1987), for example.

Extending transformation techniques to moncotyledonous plants, in particular the Gramineae is of great importance because members of this family (e.g. maize, rice, wheat and barley) constitute a substantial portion of the world's food crop. Transgenic rice plants have been developed by the introduction of foreign DNA into protoplasts by electroporation (Shimamoto *et al*, 1989). Recently, the regeneration of fertile transgenic maize plants has been achieved. Cells of embryogenic maize suspension cultures were targets for microprojectile bombardment-mediated DNA delivery (Gordan-Kamm *et al*, 1990).

The development of this technology may pave the way for the introduction of the *A.nidulans* nitrate gene cluster into the genome of a crop plant. It is likely that vectors carrying the cDNAs of these genes harnessed to plant promoter and terminator signals will have to be developed due to the differences between the transcriptional apparatus of higher plants and fungi.

In conclusion, there is much work that can be done in order to complete the picture, not only of nitrate transport but of the whole nitrate assimilatory process.

Chapter 5

References

Adams G. A. and Rose J. K., 1985. Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. Cell. **41**: 1007-1015

Amara S. G., Jones V., Rosenfeld M. G., Ong E. S. and Evans R. M., 1982. Alternative RNA splicing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature. **298**: 240

Ames G. F., 1985. The histidine transport system of *Salmonella typhimurium*. Current topics in membranes and transport. **23**: 103-119

Ames G. F. and Higgins C. F., 1983. The organisation, mechanism of action and evolution of periplasmic transport systems. TIBS. 8: 97-100

Arends H. and Sebald W., 1984. Nucleotide sequence of the cloned mRNA and gene of the ADP/ATP carrier from *Neurospora crassa*. EMBO J. 3: 377-382

Arndt K. and Fink G., 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5'TGACTC3' sequences. Proc. Natl. Acad. Sci. USA. 83: 8516-8520

Arst Jr H. N., 1984. Regulation of gene expression in *Aspergillus nidulans*. Microbiol. Sci. **1**: 137-141

Arst Jr H. N., MacDonald D. W. and Cove D. J., 1970. Molybdate metabolism in *Aspergillus nidulans* I. Mutations affecting nitrate reductase and/or xanthine dehydrogenase. Mol. Gen. Genet. **108**: 129-145

Arst Jr H. N. and Cove D. J., 1973. Nitrogen metabolite repression in *Aspergillus nidulans*. Mol. Gen. Genet. **126**: 111-141

Arst Jr H. N., Tollervey D. W. and Sealey-Lewis H. M., 1982. A possible regulatory gene for the molybdenum-containing cofactor in *Aspergillus nidulans*. J. Gen. Microbiol. **128**: 1083-1093

Arst Jr H. N., Rand K. N. and Bailey C. R., 1979. Do the tightly linked structural genes for the nitrate and nitrite reductases in *Aspergillus nidulans* form an operon? Evidence of an insertional translocation that separates them. Mol. Gen. Genet. **174**: 89-100

Back E., Burkhart W., Moyer M., Privalle L. and Rothstein S., 1989. Molecular cloning of spinach nitrite reductase. pp **284-296**. In Molecular and genetic aspects of nitrate assimilation. Ed. Wray, J.L. and Kinghorn, J.R. Oxford University Press, Oxford.

Baer B. W. and Rhoades D., 1983. Eukaryotic RNA polymerase II binds to nucleosome cores from transcribed genes. Nature. **301**: 482-488

Ballance D. J., 1986. Sequences important for gene expression in filamentous fungi. Yeast. **2**: 229-236

Barton K. A., Whiteley H. and Yang N., 1987. *Bacillus thuringiensis* δ-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to Lepidopteran insects. Plant Physiol. **85**: 1103-1109

Baum J. A., Geever R. and Giles N. H., 1987. Expression of *qa-1F* activator protein: identification of upstream binding sites in the *qa* gene cluster and localisation of the DNA-binding domain. Mol. Cell. Biol. **7**: 1256-1266

Bennetzen J. L. and Hall B. D., 1982. The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase I. J. Biol. Chem. **257**: 3018-3025

Berget S. M., 1984. Are U4 small nuclear ribonucleoproteins involved in polyadenylation? Nature. **309**: 179-182

Blake C., 1983. Exons and the evolution of proteins. TIBS. **8**: 11-13

Blobel G., 1980. Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA. **77**: 1496-1500

Boing J. T. P., 1982. Enzyme production. In Industrial microbiology. pp **634-708**. Ed. Reed, G. Macmillan, Byfleet, Surrey.

Botstein D. and Shortle D., 1985. Strategies and applications of *in vitro* mutagenesis. Science **229**: 1193-1201

Boyd D. and Beckwith J., 1990. The role of charged amino acids in the localisation of secreted and membrane proteins. Cell. **62**: 1031-1033

Bradford M. M., 1976. A rapid and sensitive method for the quantitation of protein utilising the principle of protein-binding. Anal. Biochem. **72**: 248-254

Brandl C. J. and Deber C. M., 1986. Hypothesis about the function of membrane-buried proline residues in transport proteins. Proc. Natl. Acad. Sci. USA. **83**: 917-921

Braun V. and Hantke K., 1981. Bacterial cell surface receptors. In Organisation of prokaryotic cell membranes II. pp **1-74**. Ed. Ghosh, B.K. CRC Press, Boca Raton FL.

Breathnach R. and Chambon P., 1981. Organisation and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. **50**: 349-383

Breitbart R. E., Andreadis A. and Nadal-Ginard B., 1987. Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes. Ann. Rev. Biochem,. **56**: 467-495

Brownlee A. G. and Arst Jr H. N., 1983. Nitrate uptake in *Aspergillus nidulans* and the involvement of the third gene of the nitrate assimilation gene cluster. J. Bacteriol. **155**: 1138-1146

Buchel D. E., Gronenberg B. and Muller-Hill, 1980. Sequence of the lactose permease gene. Nature. **283**: 541-545

Butler J. S., Sadhale P. P. and Platt T., 1990. RNA processing in vitro produces mature 3' ends of a variety of *Saccharomyces cerevisiae* mRNAs. Mol. Cell. Biol. **10**: 2599-2605

Caddick M. X., Arst Jr H. N., Taylor L. H., Johnson R. I. and Brownlee A. G., 1986. Cloning of the regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. EMBO J. **5**: 1087-1090

Campbell E. I., Unkles S. E., Macro J. A., van den Hondel C. A. M. J. J., Contreras R. and Kinghorn J. R., 1989. Improved transformation efficiency of *Aspergillus niger* using the homologous *niaD* gene for nitrate reductase. Curr. Genet. **16**: 53-56

Campbell W. H. and Kinghorn J. R., 1990. Functional domains of assimilatory nitrate reductases and nitrite reductases. TIBS. **15**: 315-319

Campo M. S., Spandidos D. A., Lang J. and Wilkie N. M., 1983. Transcriptional control signals in the genome of bovine papillomavirus type 1. Nature. **303**: 77-80

Cathala G., Savouret J., Mendez B., West B., Karin M., Martial J. A. and Baxter J. D., 1983. A method for isolation of intact, translationally active ribonucleic acid DNA. **2**: 329-335

Celenza J. L., Marshall-Carlson L. and Carlson M., 1988. The yeast *SNF3* gene encodes a glucose transporter homologous to the mammalian protein. Proc. Natl. Acad. Sci. USA. **85**: 2130-2134

Clarke L. and Carbon J., 1976. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E.coli* genome. Cell. **9**: 91-99

Clements J. M. and Roberts C. F., 1986. Transcription and processing signals in the 3-phosphoglycerate kinase (*PGK*) gene from *Aspergillus nidulans*. Gene. **44**: 97-105

Clutterbuck A .J., 1974. *Aspergillus nidulans*. In Handbook of genetics. pp **447-510**. Ed. King R.C. Plenum Press, New York.

Clutterbuck A. J., 1984. Loci and linkage map of the filamentous fungus *Aspergillus nidulans*. In Genetic maps. pp **265-273**. Ed. O'Brien, S.J. Cold Spring Harbour, New York.

Cohen B. L., 1972. Ammonium repression of extracellular protease in *Aspergillus nidulans*. J. Gen. Microbiol. **71**: 293-299

Coppola J. A., Field A. S. and Luse D. S., 1983. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. Proc. Natl. Acad. Sci. USA. **80**: 1251-1255

Corden J., Wasylyk B., Buchwalder A., Sassone-Corsi P., Kedinger C. and Chambon P., 1980. Promoter sequences of eukaryotic protein-coding genes. Science. **209**: 1406-1414

Corrick C. M., Twomey A. P. and Hynes M. J., 1987. The nucleotide sequence of the *amdS* gene of *Aspergillus nidulans* and the molecular characterisation of 5' mutants. Gene. **53**: 63-71

Cove D. J., 1966. The induction and repression of nitrate assimilation in the fungus *Aspergillus nidulans*. Biochim. Biophys. Acta. **113**: 51-56

Cove D. J., 1970. Control of gene action in *Aspergillus nidulans*. Proceedings of the Royal Society (London) Series B. **176**: 267-275

Cove D. J., 1976 a. Chlorate toxicity in *Aspergillus nidulans*: the selection and characterisation of chlorate resistant mutants. Heredity. **36**: 191-203

Cove D. J., 1976 b. Chlorate toxicity in *Aspergillus nidulans*. Molec. Gen. Genet. **146**: 147-159

Cove D. J., 1979. Genetic studies of nitrate assimilation in *Aspergillus nidulans*. Biol. Revs. **54**: 291-237

Cove D. J. and Pateman J. A., 1963. Independently segregating loci concerned with nitrate assimilation in *Aspergillus nidulans*. Nature. **168**: 262-263

Cove D. J. and Pateman J. A., 1969. Autoregulation of the synthesis of nitrite reductase in *Aspergillus nidulans*. J. Bacteriol. **97**: 1374-1378

Cox G. B., Webb D., Godovac-Zimmerman J. and Rosenberg H., 1988. Arg-220 of the PstA protein is required for phosphate transport through the phosphate specific transport system in *Escherichia coli* but not for alkaline phosphatase repression. J. Bacteriol. **170**: 2283-2286

Crawford N. M. and Campbell W. H., 1990. Fertile fields. The plant cell. **2**: 829-835

Cybis J. and Weglenski P., 1972. Arginase induction in *Aspergillus nidulans*. The appearance and decay of the coding capacity of messenger. Eur. J. Biochem. **30**: 262-268

Da Silva A. J. F., Whittington H., Clements J., Robert C. and Hawkins A. R., 1986. Sequence analysis and transformation by the catabolic 3-dehydroquinase (*QutE*) gene from *Aspergillus nidulans*. Biochem. J. **240**: 481-488

Dalbey R. E., 1990. Positively charged residues are important determinants of membrane protein topology. TIBS. **15**: 253-257

Darnell J. E. J., 1982. Variety in the level of gene control in eukaryotic cells. Nature. **297**: 365-371

Davis L. G., Dibner M. D. and Battey J. F., 1986. Methods in molecular biology. Elsevier, New York.

Davis M. A. and Hynes M. J., 1987. Complementation of *areA*$^-$ regulatory gene mutations of *Aspergillus nidulans* by the heterologous regulatory gene *nit-2* of *Neurospora crassa*. Proc. Natl. Acad. Sci. USA. **84**: 3753-3757

Dewey M. M. and Barr L., 1970. Some considerations about the structure of cellular membranes. Current topics in membranes and transport. **1**: 1-33

Doolittle W. F. and Sapienza C., 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature. **284**: 601-603

Dorn A., Bollenkens J., Staub A., Benoist C. and Mathis D., 1987. A multiplicity of CCAAT box binding proteins. Cell. **50**: 863-872

Dreyfruss G., 1986. Structure and function of nuclear and cytoplasmic ribonucleic particles. Ann. Rev. Cell. Biol. **2**: 459-498

Dugaiczyk A., Boyer H. W. and Goodman H. M., 1975. Ligation of *EcoRI* endonuclease-generated fragments into linear and circular structure. J. Mol. Biol. **96**: 171-184

Dynan W. S., Sazer S., Tjian R. and Chimke R. T., 1985. Transcription factor SP1 recognises a DNA sequence in the mouse dihydrofolate reductase promoter. Nature. **319**: 246-248

Dynan W. S. and Tjian R., 1985. Control of eukaryotic messenger RNA synthesis by sequence specific DNA- binding proteins. Nature. **316**: 774-778

Eisenberg D., Schwarz E., Komaromy M. and Wall R., 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. **179**: 125-142

Eisenberg D. and Hill C. P., 1989. Review on protein crystallography: more surprises ahead. TIBS. **14**: 260-264

Engelman D. M. and Steitz T. A., 1981. The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis. Cell. **23**: 411-422

Engelman D. M., Steitz T. A. and Goldman A., 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Ann. Rev. Biophys. Biophys. Chem. **15**: 321-353

Feinberg A. P. and Vogelstein B., 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**: 6-13

Ferenci T., 1989. Selectivity in solute transport: Binding sites and channel structure in maltoporin and other bacterial sugar transport proteins. Bioessays. **10**: 3-7

Fidel S., Doonan J. M. and Morris N. R., 1988. *Aspergillus nidulans* contains a single actin gene which has unique intron locations and encodes a γ-actin. Gene. **70**: 283-293

Fincham J. R. S., 1989. Transformation in fungi. Microbiol. Revs. **53**: 148-170

Firoozabady E., DeBoer D. L., Merlo D. J., Halk E. L., Amerson L. N., Rashka K. E. and Murray E. E., 1987. Transformation of cotton (*Gossypium hirstutum* L.) by regeneration of transgenic plants. Plant Mol. Biol. **10**: 105-116

Foster D. L., Boublik M. and Kaback H. R., 1983. Structure of the *lac* carrier protein of *Escherichia coli*. J. Biol. Chem. **258**: 31-34

Fu Y-H. and Marzluf G. A., 1989. Isolation of *nit-4*, the minor nitrogen regulatory gene which mediates nitrate induction in *Neurospora crassa*. J. Bacteriol. **171**: 4067-4070

Fu Y-H. and Marzluf G. A., 1990. *nit-2*, the major positive-acting nitrogen regulatory gene of *Neurospora crassa*, encodes a sequence specific DNA- binding protein. Proc. Natl. Acad. Sci. USA. **87**: 5331-5335

Furuichi Y., LaFiandra A. and Shatkin A. J., 1977. 5'-terminal structure and mRNA stability. Nature. **266**: 235-239

Garnier J., Osguthorpe D. J. and Robson B., 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**: 97-120

Geever R. F., Huiet L., Baum J. A., Tyler B. M., Patel V. B., Rutledge B. J., Case M. E. and Giles N. H., 1989. DNA sequence, organisation and regulation of the *qa* gene cluster of *Neurospora crassa*. J. Mol. Biol. **207**: 15-34

Gil A. and Proudfoot N. J., 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit ß-globin mRNA 3' end fromation. Cell. **49**: 399-406

Giles N. H., Case M. E., Baum J., Geever R., Huiet L. Patel V. and Tyler B., 1985. Gene organisation and regulation in the *qa* (Quinic acid) gene cluster of *N.crassa*. Microbiol Revs. **49**: 338-358

Glover C. V. C. I., 1989. Sequence-specific protein-DNA recognition by transcriptional regulatory proteins. Plant Molecular biology reporter. **7**: 183-208

Goc A. and Weglenski P., 1988. Regulatory region of the *Aspergillus nidulans argB* gene. Curr. Genet. **14**: 425-429

Goldberg A. L. and Dice J. F., 1974. Intracellular protein degradation in mammalian and bacterial cells. Ann. Rev. Biochem. **43**: 835-869

Gordan-Kamm W. J., Spencer T. M., Mangano M. L., Adams T. R., Daines R. J., Start W. G., O'Brien J. V., Chambers S. A., Adams J. W.R., Willets N. G., Rice T. B., Mackey C. J., Krueger R. W., Krausch A. P. and Lemaux P. G., 1990. Transformation of maize cells and regeneration of fertile plants. The Plant Cell. **2**: 603-618

Goyon C., Faugeron G. and Rossigol J., 1988. Molecular cloning and characterisation of the *met2* gene from *Ascobolus immersus*. Gene. **63**: 297-308

Green M. R., 1986. Pre-mRNA splicing. Ann. Rev. Genet. **20**: 671-708

Grieves P. A., 1990. Molecular and genetical analysis of the nitrate cluster of *Aspergillus nidulans*. Ph.D. thesis. St. Andrews. U.K.

Grieves P., Punt P. R., Montague P., Kinghorn J. K. and van den Hondel C. A. M. J. J., in press. A twin reporter system for *Aspergillus nidulans*. Gene.

Grifo J. A., Tahara S. M., Morgan M. A., Shatkin A. J. and Merrick W. C., 1983. New initiation factor required for globin mRNA translation. J. Biol. Chem. **258**: 5804-5810

Groudine M., Eisenman R. and Weintraub H., 1981. Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. Nature. **292**: 311-317

Gwynne D. I., Buxton F. P., Sibley S., Davis R. N., Lockington R. A., Scazzocchio C. and Sealey-Lewis H. M., 1987. Comparison of the *cis*-acting control regions of two coordinately controlled genes involved in ethanol utilisation in *Aspergillus nidulans*. Gene. **51**: 205-216

Hanahan D., 1983. Studies of transformation of *Escherichia coli* with plasmids, J.Mol. Biol. **166**: 557-580

Hawkins A. R., Lamb H. K., Smith M., J.W. K. and Roberts C. F., 1988. Molecular organisation of the quinic acid utilisation (*QUT*) gene cluster in *Aspergillus nidulans*. Mol. Gen. Genet. **214**: 224-231

Hengge R. and Boos W., 1983. Maltose and lactose transport in *Escherichia coli*. Examples of two different types of concentrative transport systems. Biochim. Biophys. Acta. **737**: 443-478

Hershey J. W. B., Duncan R. and Mathews M. B., 1986. Mechanisms of translational control. In Translational control. pp **1-18**. Cold Sring Harbour, New York.

Hicks J. B., Hinnen A. and Fink G. R., 1979. Properties of yeast transformation. Cold Spr. Harb. Symp. **43**: 1305-1313

Higgins C. F., Gallagher M. P., Mimmack M. C. and Pearce S. R., 1988. A family of closely related ATP-binding subunits from prokaryotic and eukaryotic cells. Bioessays. **8**: 111-116

Hoffman W., 1985. Molecular characterisation of the *CAN1* locus in *Saccharomyces cerevisiae*. J. Biol. Chem. **260**: 11831-11837

Hohn T. M. and Beremand P. D., 1989. Isolation and nucleotide sequence of a sesquiterpene cyclase gene from the trichothecene-producing fungus *Fusarium sporotrichioides*. Gene. **79**: 131-138

Holm L., 1986. Codon usage and gene expression. Nucleic Acids Res. **14**: 3075-3087

Huiet L. and Giles H. N., 1986. The *qa* repressor gene of *Neurospora crassa*: wild type and mutant nucleotide sequences. Proc. Natl. Acad. Sci. USA. **83**: 3381-3385

Hunter S., Provasoli L., Schatz A. and Haskins C. P., 1950. Some approaches to the study of the role of metals in the metabolism of micro-organisms. Proceedings of the American Philosphical Society. **94**: 152-170

Hynes M. J., 1973. The effect of lack of a carbon source of nitrate reductase activity in *Aspergillus nidulans*. J. Gen. Microbiol. **79**: 155-157

Hynes M. J., 1975. Studies on the role of the *areA* gene in the regulation of nitrogen catabolism in *Aspergillus nidulans*. Aust. J. Biol. Sci. **28**: 301-313

Jacob F. and Monod J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. **3**: 318-356

Johnston S. A., Salmeron J. M. and Dincher S. S., 1987. Interaction of positive and negative regulatory proteins in the galactose regulon of yeast. Cell. **50**: 143-146

Johnstone I. L., 1985. Transformation of *Aspergillus nidulans*. Microbiol. Sci. **2**: 307-311

Johnstone I. L., Hughes S. G. and Clutterbuck A. J., 1985. Cloning an *Aspergillus nidulans* developmental gene by transformation. EMBO J. **4**: 1307-1311

Johnstone I. L., McCabe P. C., Greaves P., Gurr S. J., Cole G. E., Brow M. A. D., Unkles S. E., Clutterbuck A. J., Kinghorn J. R. and Innis M. A., 1990. Isolation and characterisation of the *crnA - niiA - niaD* gene cluster for nitrate assimilation in *Aspergillus nidulans*. Gene. **90**: 181-192

Jones K. A. and Tjian R., 1985. Sp1 binds to promoter sequences and activates herpes simplex virus "immediate early" gene transcription in vitro. Nature. **317**: 179-182

Jones N., 1990. Transcriptional regulation by dimerisation: two sides to an incestuous relationship. Cell. **61**: 9-11

Jund R., Weber E. and Chevallier M-R., 1988. Primary structure of the uracil transport protein of *Saccharomyces cerevisiae*. Eur. J. Biochem. **171**: 417-424

Kaback H. R., 1983. The *lac* carrier protein in *Escherichia coli*. J. Membrane Biol. **76**: 95-112

Kaback H. R., Bibi E. and Roepe P. D., 1990. ß-galactosidase transport in *Escherichia coli*: a functional dissection of *lac* permease. TIBS. **15**: 355-358

Kapoor K. K., Chaudry K. and Tauro P., 1982. Citric Acid. In Industrial microbiology. pp **709-747**. Ed. Reed, G. Macmillan, Byfleet Surrey.

Keegan L., Gill G. and Ptashne M., 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science. **231**: 699-704

Khoury G. and Gruss P., 1983. Enhancer elements. Cell. **33**: 313-314

Kinghorn J. R. and Campbell E. I., 1989. Amino acid sequence relationships between bacterial, fungal and plant nitrate reductase and nitrite reductase proteins. In Molecular and genetic aspects of nitrate assimilation. pp **387-404**. Ed. Wray, J.L. Kinghorn, J.R. Oxford University Press, Oxford.

Kinnaird J. H. and Fincham J. R. S., 1983. The complete nucleotide sequence of the *N.crassa am* (NADP-specific glutamate dehydrogenase) gene. Gene. **26**: 253-260

Kolar M., Punt P. J., van den Hondel C. A. M. J. J. and Schwab H., 1988. Transformation of *Penicillium chrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion gene. Gene. **62**: 127-134

Kolata G., 1985. Fitting methylation into development. Science. **228**: 1183-1184

Kozak M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. **44**: 283-292

Kudla B., Caddick M. A., Langdon T., Martinez-Rossi N. M., Bennet C. F., Sibley S., Davies R. W. and Arst Jr H. N., 1990. The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger. EMBO J. **9**: 1355-1364

Kyte J. and Doolittle R. F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**: 105-132

Lafferty M. A. and Garrett R. H., 1974. Purification and properties of the *Neurospora crassa* assimilatory nitrite reductase. J. Biol. Chem. **249**: 7555-7567

Landschulz W. H., Johnson P. F. and McKnight S. L., 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science. **243**: 1681-1688

Langford C. J. and Gallwitz D., 1983. Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell. **33**: 519-527

Langford C., Nellen W., Neising J. and Gallwitz D., 1983. Yeast is unable to exise foreign intervening sequences from hybrid gene transcripts. Proc. Natl. Acad. Sci. USA. **80**: 1496-1560

Larsson and Ingermarsson, 1989. Molecular aspects of nitrate uptake in higher plants. In Molecular and genetic aspects of nitrate assimilation. pp **3-14**. Ed. Wray, J.L. Kinghorn, J.R. Oxford University Press, Oxford.

Leff S. E., Rosenfeld M. G. and Evans R. M., 1986. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Ann. Rev. Biochem. **55**: 1091-1117

Legerton T. L. and Yanofsky C., 1985. Cloning and characterisation of the multifunctional *his-3* gene of *Neurospora crassa*. Gene. **39**: 129-140

LeJohn H. B., 1989. Structure and expression of fungal calmodulin gene. J. Biol. Chem. **264**: 19366-19372

Lodish H. F., 1988. Transport of secretory and membrane glycoproteins from the rough endoplasmic reticulum to the Golgi; a rate limiting step in protein maturation and secretion. J. Biol. Chem. **263**: 2107-2110

MacDonald D. W. and Cove D. J., 1974. Studies on temperature sensitive mutants affecting the assimilatory nitrate reductase of *Aspergillus nidulans*. Eur. J. Biochem. **47**: 107-110

MacIntyre S., Eschbach M., Schwarz H. and Ehring R., 1989. Topological analysis of the amino teminal region of lactose permease using the *Escherichia coli* outer membrane protein as a marker. FEBS. **247**: 396-400

Macpherson A. J. S., Jones-Mortimer M. C., Horne P. and Henderson P. J. F., 1983. Identification of the GalP galactose transport protein of *Escherichia coli*. J. Biol. Chem. **258**: 4390-4396

McLauchlan J., Gaffney D., Whitton J. L. and Clements J. B., 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucl. Acids. Res. **13**: 1347-1368

McNeil J. B. and Smith M., 1985. *Saccharomyces cerevisiae* mRNA 5'-end positioning: analysis by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base pairs. Mol. Cell. Biol. **5**: 3545-3551

Ma J. and Ptashne M., 1987. The carboxy terminal 30 amino acids of GAL4 are recognised by GAL80. Cell. **50**: 137-142

Maiden M. C. J., Davis E. O., Baldwin S. A., Moore D. C. M. and Henderson P. J. F., 1987. Mammalian and bacterial sugar transport proteins are homologous. Nature. **325**: 641-643

Mann B. J., Bowmann B. J., Grotelueschem J. and Metzenberg R. L., 1989. Nucleotide sequence of *pho-4*$^+$, encoding a phosphate repressible phosphate permease of *Neurospora crassa*. Gene. **83**: 281-289

Manoil C. and Beckwith J., 1986. A genetic approach to analysing membrane protein topology. Science. **233**: 1403-1408

Marzluf G. A., 1970. Genetic and metabolic controls for sulfate metabolism in *Neurospora crassa*: Isolation and study of chromate resistant and sulfate transport negative mutants. J. Bacteriol. **102**: 716-721

Mercola M., Wang X., Olsen J. and Calame K., 1983. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. Science. **221**: 663-665

Mermelstein F. H., Flores O. and Reinberg D., 1989. Initiation of transcription by RNA polymerase II Biochim. Biophys. Acta. **1009**: 1-10

Miller J., McLachlan A. D. and Klug A., 1985. Repetitive zinc binding domains in the protein transcription factor III A from *Xenopus* oocytes. EMBO J. **4**: 1609-1614

Mills F. C., Fisher L. M., Kuroda R., Ferd A. M. and Gould H. J., 1983. DNase I hypersensitive sites in the chromatin of human μ immunoglobulin heavy-chain genes. Nature. **306**: 809-812

Mitchell P. and Tjian R., 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science. **245**: 371-378

Mueckler, Caruso C., Baldwin S. A., Panico M., Blench I., Morris H. R., Allard W. J., Lienhard G. E. and Lodish H. F., 1985. Sequence and structure of a human glucose transporter. Science. **229**: 941-945

Mullaney E. J., Hamer J. E., Roberti K. A., Yelton M. M. and Timberlake W. E., 1985. Primary structure of the *trpC* gene from *Aspergillus nidulans*. Molec. Gen. Genet. **199**: 37-45

Murre C., McCaw P. S. and Baltimore D., 1989. A new DNA-binding and dimerisation motif in immunoglobulin ehhancer binding daughterless, MyoD and myc proteins. Cell. **56**: 777-783

Nakao T., Yamato I. and Anraku Y., 1987. Nucleotide sequence of *putP*, the proline carrier gene of *Escherichia coli*. Mol. Gen. Genet. **208**: 70-75

Nehlin J. O., Carlberg M. and Ronne H., 1989. Yeast galactose permease is related to yeast and mammalian glucose transporters. Gene. **85**: 313-319

Newbury S. F., Glazebrook and Radford A., 1986. Sequence analysis of the *pyr-4* (orotidine 5'-P decaroxylase) gene of *Neurospora crassa*. Gene. **43**: 51-58

Nilsson I. and von Heijne G., 1990. Fine-tuning of the topology of a polytopic membrane protein: Role of positively and negatively charged amino acids. Cell. **62**: 1135-1141

Nordheim A., Pardue M. L., Lafer E. M., Moller A., Stollar B. D. and Rich A., 1981. Antibodies to left-handed Z-DNA bind to interband regions of *Drosophila* polytene chromosomes. Nature. **294**: 417-422

Omata T., 1990. Cloning, nucleotide sequence and mutagenesis of the *nrtA* gene encoding the 45-kD protein involved in nitrate transport of the cyanobacterium*Synechoccus* PCC 7942. Nitrate assimilation: molecular and genetic aspects. 3rd international symposium, Bombannes, France.

Pateman J. A., Cove D. J., Rever B. M. and Roberts D. B., 1964. A common co-factor for nitrate reductase and xanthine dehydrogenase which also regulates the synthesis of nitrite reductase. Nature. **201**: 58-60

Pateman J. A., Rever B. M. and Cove D. J., 1967. Genetic and biochemical studies of nitrate reductase in *Aspergillus nidulans*. Biochem. J. **104**: 103-111

Pateman J. A., Doy C. H., Olsen J. E., Norris U., Creaser E. H. and Hynes M., 1983. Regulation of alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (AldDH) in *Aspergillus nidulans*. Proc. R. Soc. B. **217**: 243-264

Paulson J. C., 1989. Glycoproteins: what are the sugar chains for? TIBS. **14**: 272-276

Pederson T., 1983. Nuclear RNA-protein interactions and messenger RNA processing . J. Cell. Biol. **97**: 1321-1326

Pfeifer K., Prezant T. and Guarente L., 1987. HAP1 activator binds to two specific upstream activation sites of different sequences. Cell. **49**: 19-27

Pickett M., Gwynne D. I., Buxton F. P. E. R., Davies R. W., Lockington R. A., Scazzocchio C. and Sealey-Lewis H. M., 1987. Cloning and characterisation of the *aldA* gene of *Aspergillus nidulans*. Gene. **51**: 217-226

Pinkert C. A., Ornitz D. M., Brinster R. L. and Palmiter R. D., 1987. An albumin enhancer located 10 kb upstream functions along with its promoter to diet efficient liver-specific expression in transgenic mice. Genes and development. **1**: 268

Pombeiro S. R. C., Martinez-Rossi N. M. and Rossi A., 1983. Nitrite toxicity in *Aspergillus nidulans*: a new locus in a *proA*1 *pabaA*6 *yA*2 strain. Genet. Res. **41**: 203-207

Pontecorvo G., Roper J. A., Hemmons D. W., McDonald K. d. and Bufton A. W., 1953. The genetics of *Aspergillus nidulans*. Adv. Genet. **5**: 141-238

Postma P. W. and Roseman S., 1976. The bacterial phosphoenolpyruvate: sugar phosphotransferase system. Biochim. Biophys. Acta. **457**: 213-257

Ptashne M., 1986. Gene regulation by proteins acting nearby and at a distance. Nature. **322**: 697-701

Ptashne M., 1988. How eukaryotic transription factors work. Nature. **335**: 683-689

Pua E., Mehra P. A., Nagy F. and Chua N., 1987. Transgenic plants of *Brassica napus* L. Bio/Technology. **5**: 815-817

Punt P. J., Dingemanse M. A., Jacobs-Meijsing B. J. M., Pouwels P. H. and van den Hondel C. A. M. J. J., 1988. Isolation and characterisation of the glyceraldehyde-3-phosphate dehydrogenase gene of *Aspergillus nidulans*. Gene. **69**: 49-57

Punt P. J., Dingemanse M. A., Kuyvenhoven A., Soede R. D. M., Pouwels P. H. and van den Hondel C. A. M. J. J., 1990. Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene-encoding glyceraldehyde-3-phosphate dehydrogenase. Gene. **93**: 111-117

Rai R., Genbauffer F. S. and Cooper T. G., 1988. Structure and transcription of the allantoate permease gene (*DAL5*) from *Saccharomyces cerevisiae*. J. Bacteriol. **170**: 266-271

Rambosek J. and Leach J., 1987. Recombinant DNA in filamentous fungi: progress and prospects. CRC Critical reviews in Biotechnology. **6**: 357-393

Rand K. N. and Arst Jr H. N., 1977. A mutation in *Aspergillus nidulans* which affects the regulation of nitrite reductase and is tightly linked to its structural genes. Mol. Gen. Genet. **155**: 67-75

Rand K. N. and Arst Jr H. N., 1978. Mutations in *nirA* gene of *Aspergillus nidulans* and nitrogen metabolism. Nature. **272**: 732-734

Rapoport T. A., 1990. Protein transport across the ER membrane. TIBS. **15**: 355-358

Redinbaugh M. G. and Campbell W. H., 1985. Quaternary structure and composition of squash NADH:nitrate reductase J. Biol. Chem. **260**: 3380-3385

Rhoads R. E., 1985. The cap structure of eukaryotic messenger RNA and its interaction with cap-binding protein. Prog. Mol. Subcell. Biol. **9**: 104

Rickenberg H. V., Cohen G. N. and Monod J., 1956. La Galactoside-permease d'*Escherichia coli* (abstract) Ann. Inst. Pasteur Paris. **91**: 829-857

Riordan J. R., Rommens J. M., Kerem B., Alon N., Rozmahel R., Grzelcak Z., Zielenski J., Lok S., Plavsic N., Chou J., Drumm M. L., Iannuzzi M. C., Collins F. S. and Tsui L., 1989. Identification of the yeast cystic fibrosis gene:-cloning and characterisation of complementary cDNA. Science. **245**: 1066-1073

Rothman and Lodish, 1977. Synchronised transmembrane insertion and glycosylation of a nascent membrane protein. Nature. **269**: 775-780

Rutledge B. J., 1984. Molecular characterisation of the *qa-4* gene of *Neurospora crassa*. Gene. **32**: 275-287

Sambrook J., Fritsch E. F. and Maniatis T., 1989. Molecular cloning: a laboratory manual. Cold Spring Harbour, New York.

Sanno Y., Wilson T. H. and Lin E. C. C., 1968. Control of permeation to glycerol in cells of *Escherichia coli*. Biochim. Biophys. Res. Commun. **32**: 344-349

Sassone-Corsi P. and Borrelli E., 1986. Transcriptional regulation by *trans*-acting factors. Trends in Genetics. **2**: 215

Scazzocchio C. and Arst Jr H. N., 1989. Regulation of nitrate assimilation in *Aspergillus nidulans*. In Molecular and genetic aspects of nitrate assimilation. pp **299-313**. Ed. Wray, J.L. Kinghorn, J.R. Oxford University Press, Oxford.

Schechtman M. G. and Yanofsky C., 1983. Structure of the trifunctional *trp-1* gene from *Neurospora crassa* and its aberrant expression in *Escherichia coli*. J. Mol. Appl. Genet. **2**: 83-99

Shapiro D. J., Blume J. E. and Nielson D. A., 1988. Regulation of messenger RNA stability in eukaryotic cells. Bioessays. **6**: 221-226

Sharp P. A., 1985. On the origin of RNA splicing and introns. Cell. **42**: 397-400

Shimamoto K., Terada R., Izawa T. and Fujimoto H., 1989. Fertile transgenic rice plants regenerated from transformed protoplasts. Nature. **338**: 274-276

Sigler P. B., 1988. Acid blobs and negative noodles Nature. **333**: 210-212

Silhavy T. J. and Beckwith J. R., 1985. Uses of *lac* fusions for the study of biological problems. Microbiol. Revs. **49**: 398-418

Singer R. H. and Penman S., 1973. Messenger RNA in HeLA cells: kinetics of formation and decay. J. Mol. Biol. **78**: 321-334

Smith C. W. J., Patton J. G. and Nadal-ginard B., 1989. Alternative splicing in the control of expression. Ann. Rev. Genet. **23**: 527-577

Soliday C. L., Dickman M. B. and Kolattukudy P. E., 1989. Structure of the cutinase gene and detection of promoter activity in the 5' flanking region by fungal transformation. J.Bacteriol. **171**: 1942-1951

Sophianopoulou V. and Scazzocchio C., 1989. The proline transport protein of *Aspergillus nidulans* is very similar to amino acid transporters of *Saccharomyces cerevisiae*. Molecular. Microbiol. **3**: 705-714

Sorger G. J., Brown D., Farzannejad M., Guerra A., Jonathon M., Knight S. and Shorda R., 1989. Isolation of a gene that down regulates nitrate assimilation and influences another regulatory gene in the same system. Mol. Cell. Biol. **9**: 4113-4117

Steiner D. F. and Oyer P. E., 1967. The biosynthesis of insulin and a probable precursor of insulin by a human islet cell ademona. Proc. Natl. Acad. Sci. USA. **57**: 473-480

Struhl K., 1989. Helix-turn-helix, zinc finger and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. TIBS. **14**: 137-140

Tanaka J-I. and Fink G. R., 1985. The histidine permease (HIP1) of *Saccharomyces cerevisiae*. Gene. **38**: 205-214

Timberlake W. E., 1978. Low repetitive DNA content in *Aspergillus nidulans*. Science. **202**: 973-975

Tognoni A., Cattaneo R., Serfling E. and Schaffner W., 1985. A novel expression selection approach allows precise mapping of the hepatitis B virus enhancer. Nucleic Acids Research. **13**: 7457-7472

Tollervey D. W. and Arst Jr H. N., 1981. Mutations to constitutivity and derepression are separate and separable in a regulatory gene of *Aspergillus nidulans*. Curr. Genet. **4**: 63-68

Tomsett A. B. and Cove D. J., 1979. Deletion mapping of the *niiA-niaD* gene region of *Aspergillus nidulans*. Genetical. Res. **34**: 19-32

Turner R. and Tjian R., 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Science. **243**: 1689-1694

Unkles S. E., Hawker K. L., Grieve C., Campbell E., Montague P. and Kinghorn J. K., 1991. The *crnA* gene encodes a nitrate permease in *Aspergilllus nidulans*. Proc. Natl. Acad. Sci. USA. **88**: 204-208

van Gorcom R. F. M., Punt P. J., Pouwels P. H. and van den Hondel C. A. M. J. J., 1986. A system for the analysis of expression signals in *Aspergillus*. Gene. **48**: 211-217

van Vunakis H. and Herriot R. M., 1956. Structural changes associated with the conversion of pepsinogen to pepsin. Biochim. Biophys. Acta. **22**: 537-543

von Heijne G. and Gavel Y., 1988. Topogenic signals in internal membrane proteins. Eur. J. Biochem. **174**: 671-678

Verner K. and Schatz G., 1988. Protein translocation across membranes. Science. **41**: 1307-1313

Villeponteau B., Lundrell M. and Martinson H., 1984. Torsional stress promotes the DNase 1 sensitivity of active genes. Cell. **39**: 469-478

Yazyu H., Shiota-Niiya S., Shimamoto T., Kanazawa F. and Tsuchiya T., 1984. Nucleotide sequence of the *melB* gene and characteristics of deduced amino acid sequence of the melibiose carrier in *Escherichia coli*. J. Biol. chem. **259**: 4320-4326

Zabarovski E. R. and Turina O. V., 1988. Rapid isolation of λ phage DNA in micro and macro variants. Nucl. Acids. Res. **16**: 10925

Zaret K. S. and Sherman F., 1982. DNA sequence required for efficient transcription termination in yeast. Cell. **28**: 563-573

Zhang R. G., Joachimiak A., Lawson C. L., Schevitz R. W., Otwinowski Z. and Sigler P. B., 1987. The crystal structure of *trp* aporepressor at 1.8 A shows how binding of tryptophan enhances DNA affinity. Nature. **327**: 591-597

Vogel H., Wright T. K. and Jahnig F., 1985. The structure of the lactose permease from Raman spectroscopy and prediction methods. EMBO J. **4**: 3625-3631

Walker J. E., Soraste M., Runswick M. J. and Gay N. J., 1982. Distantly related sequences in the α and ß-subunits of ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**: 945-951

Walter P., Gilmore R. and Blobel G., 1984. Protein translocation across the endoplasmic reticulum. Cell. **38**: 5-8

Wang X. and Calame K., 1986. SV40 enhancer-binding factors are required at the establishment but not the maintenance step of enhancer-dependent transcriptional activation. Cell. **47**: 241-247

Ward M. and Turner G., 1986. The ATP synthase subunit 9 gene of *Aspergillus nidulans* sequence and transcription. Mol. Gen. Genet. **205**: 331-338

Weber E., Rodriguez C., Chevalier M. R. and Jund R., 1990. The purine-cytosine permease gene of *Saccharomyces cerevisiae*: primary structure and deduced protein sequence of the *FCY2* gene product. Molecular Microbiol. 4: 585-596

Weintraub H., 1985. Assembly and propagation of repressed and derepressed chromosomal states. Cell. **42**: 705-711

Weisbrod S., 1982. Active chromatin. Nature. **284**: 601-603

Wessels H. P. and Spiess M., 1988. Insertion of multispanning membrane proteins occurs sequentially and requires only one signal sequence. Cell. **55**: 61-70

Wickens M., 1990. How the messenger got its tail: addition of poly(A) in the nucleus. TIBS. **15**: 277-281

Wickner W., 1979. The assembly of proteins into and across membranes: the membrane trigger hypothesis. Ann. Rev. Biochem. **48**: 23-45

Wickner W., 1989. Secretion and membrane assembly. TIBS. **14**: 280-283

Wray J. L., 1989. Towards the molecular characterisation of nitrate uptake in higher plants. Aspects of applied biology. **22**: 357-364

Yao B., Sollitti P. and Marmur J., 1989. Primary structure of the maltose-permease-encoding gene of *Saccharomyces carlsbergensis*. Gene. **79**: 189-197